

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 13:57:33 ; Search time 12.98 Seconds

(without alignments)
718,843 Million cell updates/sec

Title: US-09-995-917A-1

Perfect score: 2027
Sequence: 1 MGMPFIEGTISFVKPHRSDS.....YVEFKRGLLEIETKFLFD 382Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	655	32.3	472	2	US-08-622-166A-2
2	655	32.3	472	2	US-08-622-166A-4
3	294.5	14.5	492	3	US-08-724-466B-2
4	294.5	14.5	492	4	US-08-882-164D-2
5	292.5	14.4	497	4	US-08-882-164D-32
6	269.5	13.3	497	3	US-08-724-466B-4
7	269.5	13.3	497	4	US-08-882-164D-4
8	178.5	8.8	510	3	US-08-606-505B-66
9	178.5	8.8	510	4	US-09-616-990-66
10	172	8.5	508	4	US-08-991-677-2
11	172	8.5	531	4	US-09-380-420C-2
12	170.5	8.4	495	1	US-08-532-065B-2
13	169.5	8.4	506	1	US-08-313-075A-38
14	169.5	8.4	506	3	US-08-606-505B-65
15	169.5	8.4	506	4	US-09-616-990-65
16	164.5	8.1	496	1	US-08-913-075A-50
17	159	7.8	509	3	US-08-948-564-18
18	159	7.8	521	3	US-08-948-564-14
19	158.5	7.8	490	1	US-08-201-118-7
20	158.5	7.8	490	2	US-08-238-821B-7
21	158.5	7.8	490	2	US-08-238-821B-7
22	156.5	7.7	496	4	PCT-US95-05744-7
23	156	7.7	507	1	US-08-457-274A-23
24	156	7.7	507	5	PCT-US95-05758-23
25	153	7.5	516	3	US-08-948-564-12
26	150	7.4	490	1	US-08-201-118-11
27	150	7.4	490	2	US-08-238-821B-11

28	150	7.4	490	5	PCT-US95-05744-11	Sequence 11, App1
29	149	7.4	490	1	US-08-201-118-5	Sequence 5, App1
30	149	7.4	490	2	US-08-238-821B-5	Sequence 5, App1
31	149	7.4	490	5	PCT-US95-05744-5	Sequence 5, App1
32	147	7.3	507	1	US-08-457-274A-22	Sequence 22, App1
33	147	7.3	507	1	PCT-US95-05758-22	Sequence 22, App1
34	146.5	7.2	496	3	US-08-881-784-1	Sequence 1, App1
35	146.5	7.2	496	4	US-09-292-768-2	Sequence 2, App1
36	146.5	7.2	496	4	US-09-292-768-64	Sequence 64, App1
37	146.5	7.2	496	4	US-09-292-768-66	Sequence 66, App1
38	144.5	7.1	513	3	US-08-948-564-6	Sequence 6, App1
39	144.5	7.1	520	2	US-09-091-432-2	Sequence 2, App1
40	143.5	7.1	511	4	US-08-991-677-4	Sequence 4, App1
41	141.5	7.0	476	1	US-08-313-075A-30	Sequence 30, App1
42	141	7.0	490	1	US-08-201-118-3	Sequence 3, App1
43	141	7.0	490	1	US-08-201-118-9	Sequence 9, App1
44	141	7.0	490	2	US-08-238-821B-3	Sequence 3, App1
45	141	7.0	490	2	US-08-238-821B-9	Sequence 9, App1

ALIGNMENTS

RESULT 1
US-08-622-166A-2
; Sequence 2, Application US/08622166A
; Patent No. 5952545
; GENERAL INFORMATION:
; APPLICANT: KONCZ, CSABA
; APPLICANT: MATYUR, JAIDEEP
; APPLICANT: SEKERES, MIKLOS
; APPLICANT: ALTMANN, THOMAS
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING CYTOCHROME
; TITLE OF INVENTION: P450-TYPE PROTEINS INVOLVED IN THE BRASSINOSTEROID SYNTHESI
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/622,166A
; FILING DATE: 27-MAR-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 0147-0153P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ. ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-622-166A-2

Query Match 32.3%; Score 655; DB 2; Length 472;
Best Local Similarity 34.5%; Pred. No. 9, 2e-60;
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;
QY 1 MGMPFIEGTISFVKPHRSDSIGTFLOORVSRGKVRKNSICGKAVVSCDDELNMFILN 60

Db 36 LGPLIGETQOLIGAVKTEPEPEPDERVARVGSVFMTHLEGEPTIFSADPETNRFLVN 95
QY 61 EGGLEFTSDYPRAMDIIGKSLLATGEIHRKLNVIISFNLKSPDPDLHCENLIS 120
Db 96 EGGLEFTSDYPRAMDIIGKSLLATGEIHRKLNVIISFNLKSPDPDLHCENLIS 155
QY 121 ILKSMKRCREVEFEKPKMFTLSVYNQSLSIKPEDPARLYLDDFLSYMGFISLPIPL 180
Db 156 NLDSMSS--RYLMEAKKTIFFELTVKQMSFDGEMSE--SLKREYLLVIGFISLPIPL 212
QY 181 PGCTYNAIKVRSNRNIHQNAIT-----EDMNAIREDDELDSITSNED----- 224
Db 213 FSTYRAIKQAR--RKVAEALTVMKRRBEEGAEKCKMMLAALLAADDGFSDEEIVD 270
QY 225 -----EEHAIRAKKGDELINMEDYCKME 249
Db 271 FLVALVAGYETSTIMTLAVKFLTEPTPLALQLEKEHEKIRAKMSDYSLEMSDYSMP 330
QY 250 FTQCVISEALRCGNIVKTVHRKATHDIKFEKYYIPKGMKVFPIPTAVHLDPSLHENPFEF 309
Db 331 FTQCVVNETLRVANIIGVRRAMTDVEIKGYKIPKGMKVFSSRAVHLDPNHKKDARTF 390
QY 310 NPMRWMT-----TAPGGGVRCVCGEGELGKQIAFFLHHLVLSYRM 350
Db 391 NPMRWMSNVYTGPSNVTFEPGGPRLCPGYELARVALSVFLHRLVYGFWSM 441

RESULT 2
US-08-622-166A-4
Sequence 4, Application US/08622166A
Patent No. 5952545
GENERAL INFORMATION:
APPLICANT: KONCZ, CSABA
APPLICANT: MATHUR, JAIDEEP
APPLICANT: SZEKERES, MIKLOS
APPLICANT: ALTMANN, THOMAS
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING CYTOCHROME
TITLE OF INVENTION: P450-TYPE PROTEINS INVOLVED IN THE BRASSINOSTEROID SYNTHESIS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/622,166A
FILING DATE: 27-MAR-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 0147-0153P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-622-166A-4

Query Match 32.3%; Score 655; DB 2; Length 472;
Best Local Similarity 34.5%; Pred. No. 9,2e-60;
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

QY 1 MGWPIGETTSFEKPHNSDSIGFLQORVSRVYKVSNCIGKAVVSCDQELMFTLQN 60
Db 36 LGPLIGETQOLIGAVKTEPEPEPDERVARVGSVFMTHLEGEPTIFSADPETNRFLVN 95
QY 61 EGGLEFTSDYPRAMDIIGKSLLATGEIHRKLNVIISFNLKSPDPDLHCENLIS 120
Db 96 EGGLEFTSDYPRAMDIIGKSLLATGEIHRKLNVIISFNLKSPDPDLHCENLIS 155
QY 121 ILKSMKRCREVEFEKPKMFTLSVYNQSLSIKPEDPARLYLDDFLSYMGFISLPIPL 180
Db 156 NLDSMSS--RYLMEAKKTIFFELTVKQMSFDGEMSE--SLKREYLLVIGFISLPIPL 212
QY 181 PGCTYNAIKVRSNRNIHQNAIT-----EDMNAIREDDELDSITSNED----- 224
Db 213 FSTYRAIKQAR--RKVAEALTVMKRRBEEGAEKCKMMLAALLAADDGFSDEEIVD 270
QY 225 -----EEHAIRAKKGDELINMEDYCKME 249
Db 271 FLVALVAGYETSTIMTLAVKFLTEPTPLALQLEKEHEKIRAKMSDYSLEMSDYSMP 330
QY 250 FTQCVISEALRCGNIVKTVHRKATHDIKFEKYYIPKGMKVFPIPTAVHLDPSLHENPFEF 309
Db 331 FTQCVVNETLRVANIIGVRRAMTDVEIKGYKIPKGMKVFSSRAVHLDPNHKKDARTF 390
QY 310 NPMRWMT-----TAPGGGVRCVCGEGELGKQIAFFLHHLVLSYRM 350
Db 391 NPMRWMSNVYTGPSNVTFEPGGPRLCPGYELARVALSVFLHRLVYGFWSM 441

RESULT 3
US-08-724-466B-2
Sequence 2, Application US/08724466B
Patent No. 6063606
GENERAL INFORMATION:
APPLICANT: Petkovich, P. Martin, White, Jay A.
APPLICANT: Beckett, Barbara R., Jones, Glenville
TITLE OF INVENTION: Retinoid Metabolizing Protein
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Blake, Cassels & Graydon
STREET: Box 25, Commerce Court West
CITY: Toronto
ZIP: M5L 1A9
COUNTRY: Canada
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
COMPUTER: COMPAQ, IBM PC compatible
OPERATING SYSTEM: MS-DOS 5.1
SOFTWARE: WORD PERFECT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,466B
FILING DATE: October 1, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/667,546
FILING DATE: June 21, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunt, John C.
REGISTRATION NUMBER: 36,424
REFERENCE/DOCKET NUMBER: 50767/00004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 863-4344
TELEFAX: (416) 863-2653
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 492 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-724-466B-2

ADDRESS: Blake, Cassels & Graydon
STREET: Box 25, Commerce Court West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5L 1A9
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
COMPUTER: COMPAQ, IBM PC compatible
OPERATING SYSTEM: MS-DOS 5.1
SOFTWARE: WORD PERFECT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/882,164D
FILING DATE: June 25, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/667,546
FILING DATE: June 21, 1996
APPLICATION NUMBER: 08/724,466
FILING DATE: October 1, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunt, John C.
REGISTRATION NUMBER: 36,424
REFERENCE/DOCKET NUMBER: 50767/00010
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 863-4344
TELEFAX: (416) 863-2653
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-882-164D-4

Query Match 13.3%; Score 269.5; DB 4; Length 497;
Best Local Similarity 23.2%; Pred. No. 1.4e-19;
Matches 100; Conservative 64; Mismatches 176; Indels 91; Gaps 13;

QY 1 MGPMPGEGTISFFKPHRSDSIGTFLOQVSRGKVFKNICGKAVVSCDOELMFIION 60
DB 49 MGPMPGEGTISFFKPHRSDSIGTFLOQVSRGKVFKNICGKAVVSCDOELMFIION 103
QY 61 ECKLFTSDYPRKAMHDLGKYSLLATGCEIHRKLVNIIISFINLTKSKPDELHC-----AE 115
DB 104 DDLRLVSVHMPASVRYTLGSGCLSLNLDSSHKKQKKVIMAFS-----RALECYVITE 158
QY 116 NLISILSKMKNC--REVEFHKEVKMFTLSVNVNQLSTIKP---EDPARLYVLODFLSY 169
DB 159 EVG-SLEQMLSCGERGLLVYPRVKRLMFRIMRILLGCEPOLAGDGDSEQLVFEEM 217
QY 170 MKGFISLPRLPGCTGTNAIKVRS--NRNINON-----AIIIE 204
DB 218 TRMLFSLPDIYPRGKGMKARNLHARLEQNIIRAKICGLRASGAGCKDALQLLIEH 277
QY 205 -----DMNNAIRRE-----DFLDSTIS-----NEDEHAIR 231
DB 278 SMRGERLDM-QALKOSSTFLLEGHETASATSLITYGLYPLVQKRELSKGLL 336
QY 232 AKKGDELLMVEDYQKMEFTQCVISEALRCGNIVTVHRAKTHDIFKEVYIPKMKVPP 291
DB 337 CKSNODMKLDMELDEQKYIGVYKFTLRNLPVPGGFVALKTFELNGYQIPKGMNVIT 396
QY 292 IFTAVHLDPSLHNPFFENMRT-----KTTAFGGGVYVCGGGLGLOLAFL 341
DB 397 SICDTHVAIEFTNKEEFNDRESAPHPEDASRFSFIPEGGLRSCVGKEFAKILKIFT 456
QY 342 HHLVLSYRKMI 352
DB 457 VELARHCDMOL 467

RESULT 8

US-08-606-505B-66
Sequence 66, Application US/08606505B
Patent No. 6114601
GENERAL INFORMATION:
APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: KIKUCHI, Shigeto
APPLICANT: SHIMADA, Yukihisa
APPLICANT: OHBAYASHI, Masaya
APPLICANT: SHIMADA, Ritsuko
APPLICANT: OKINAKA, Yasushi
TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
COMPUTER: IBM PS/7
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/606,505B
FILING DATE: 23-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44965/92
FILING DATE: 02-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Perry, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 510 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Eustoma russellianum
FEATURE:
NAME/KEY: CDS
LOCATION: 92 to 1621
IDENTIFICATION METHOD: by experiment
US-08-606-505B-66

Query Match 8.8%; Score 178.5; DB 3; Length 510;
Best Local Similarity 20.5%; Pred. No. 4.5e-10;
Matches 95; Conservative 68; Mismatches 160; Indels 141; Gaps 21;

QY 1 MGPMPGEGTISFFKPHRSDSIGTFLOQVSRGKVF--FKSNICGKAVVSCDOELMFI 57
DB 42 IGMPIVGLARLGLTMRH-----VALANNAKKYGPVMTYKVGSC-GLAVASTPEAKAR- 93
QY 58 LONEGKLTSDYPRK--AMHDLGKYSLLA---TGEIHRKLVNIIISFINLTKSKPDEL 111
DB 94 LKTLDMFNSRPPNGATHLAVNAODMVFADYGPBKKLRLKLSNIHI----- 140
QY 112 HCAENLSTILSKMKCRVEF-----H-----KEVKFTLSVNVNQLSTIK 153
DB 141 -----LGKALQGWEEVRKKELGYLMYMAESGRHQPVVSEMLTYAMANNLGOVMLSK 195
QY 154 PEDPARLYVLO-----DF-----LSYMKGFISLPRLPGCTGTNAIKVRSN-RNTHQ-- 199
DB 196 -----RVFGSGSESNFEDMNVELMTVAGYFNIGDFISIAMMDLOGIGCKRLHKKF 250
QY 200 -----NAIIEDMNNAIRRE-----DFLDSTISNEDEHA-----INA-----KKCD 236

Db 251 DALITRLLEEHASAEHRKSGPDLDFVYVANGNNSGERRQTVNITKALLINMTAGTDTIS 310
QY 237 GELLAN-----EDYQKMEFTQCVISEALR-CGNI 264
Db 311 SSVIEWALAEELKNPILIRRAOEMDVGIGRDRFLEADISKLPLYQAIKKEAFRKHPSY 370
QY 265 VKTVHRRKATHDIKFEKVEYIPKGMKVPFIFTAVHLDPSLHNEPFEFNPMT----- 315
Db 371 PLNLPRIASQACEVNGHYIPKGRILSVNITWALGRDPSVWENPNEFNPDRFLERKNAKIDP 430
QY 316 -----KTAFGGAVRCPGEGELKQIAEFLHNLVLSYRKIKS 354
Db 431 RGNDFELIPFAGARRICAGTRIGILVETLGTIVHSFWELPS 474

RESULT 9
US-09-616-990-66
Sequence 66, Application US/09616990
Patent No. 6232109
GENERAL INFORMATION:
APPLICANT: KIKUCHI, Yasuhiro
KIYOKAWA, Shigeto
SHIMADA, Yukinisa
OHBAIYASHI, Masaya
SHIMADA, Ritsuko
OKINAKA, Yasushi
TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
COMPUTER: IBM PS/4
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/616,990
FILING DATE: 14-Jul-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Percy, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 510 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Eustoma russellianum
FEATURE:
NAME/KEY: CDS
LOCATION: 92 to 1621
IDENTIFICATION METHOD: by experiment
SEQUENCE DESCRIPTION: SEQ ID NO: 66
US-09-616-990-66

Query Match 8.8%; Score 178.5; DB 4; Length 510;
Best Local Similarity 20.5%; Pred. No. 4.5e-10;
Matches 95; Conservative 68; Mismatches 160; Indels 141; Gaps 21;

QY 1 MGPFFICE-TISFEPKRRSDSIGTFLQQRVSRGKY--FKSNIGCGKAVVSCDQELNMF 57
Db 42 IGWPVLGALRLIGTMRP-----VALANNAKKYGPWYILKVSQ-CGLAASPEAKAF- 93
QY 58 IONEGKLFSTDVPK--AMHDILGKYSLLA---TGEHRRKKNVYISINILTKSPDPL 111
Db 94 LKIDMNFESNRRPNACATILATNAQDMVFADYGPRKRLKLSNHI----- 140
QY 112 HCAENLISISILSKMKCREVEF-----H-----KEVMKFTLSVAVNQLLSIK 153
Db 141 -----LGKRALQGEVEVRKKEIGYMLYAAESGRHQPVVSEMLTYAANNLGGYMLSK 195
QY 154 PEDRALYVLQ-----DF-----LSYMGFTSLPIPLPGTGTNAIKVRSN-RNIHQ-- 199
Db 196 -----RVFGSQSESNEFDMVVELMTVAGYFNIGDFIPSIAMMDQIGGCMKRILKKF 250
QY 200 -----NAIEDMNNALREE-----DFLDSIISNEDEEHA-----TRA-----KKGD 236
Db 251 DALITRLLEEHASAEHRKSGPDLDFVYVANGNNSGERRQTVNITKALLINMTAGTDTIS 310
QY 237 GELLAN-----EDYQKMEFTQCVISEALR-CGNI 264
Db 311 SSVIEWALAEELKNPILIRRAOEMDVGIGRDRFLEADISKLPLYQAIKKEAFRKHPSY 370
QY 265 VKTVHRRKATHDIKFEKVEYIPKGMKVPFIFTAVHLDPSLHNEPFEFNPMT----- 315
Db 371 PLNLPRIASQACEVNGHYIPKGRILSVNITWALGRDPSVWENPNEFNPDRFLERKNAKIDP 430
QY 316 -----KTAFGGAVRCPGEGELKQIAEFLHNLVLSYRKIKS 354
Db 431 RGNDFELIPFAGARRICAGTRIGILVETLGTIVHSFWELPS 474

RESULT 10
US-08-991-677-2
Sequence 2, Application US/08991677A
Patent No. 6252135
GENERAL INFORMATION:
APPLICANT: Chiang, Vincent L
APPLICANT: Carttaway, Daniel T
APPLICANT: Smeltzer, Richard H
TITLE OF INVENTION: Production of syringyl lignin in gymnosperms
FILE REFERENCE: 50617
CURRENT APPLICATION NUMBER: US/08/991,677A
EARLIER FILING DATE: 1997-12-16
EARLIER APPLICATION NUMBER: US 60/033,381
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 508
TYPE: PRT
ORGANISM: Liquidambar styraciflua
US-08-991-677-2

Query Match 8.5%; Score 172; DB 4; Length 508;
Best Local Similarity 25.4%; Pred. No. 2.1e-09;
Matches 52; Conservative 31; Mismatches 60; Indels 62; Gaps 7;

QY 186 TNAIKVR-----SNRNHONALIEDMNNALREDFLDSIISNEDEEHAIRAKKGDGE 238
Db 303 TTAISVEMAMAEILKNRPAVQKA-----QEELDNVIGSF-----R 337
QY 239 LTNMEDYQKMEFTQCVISEALRCGNIYKVH-----RKATHDIKFEKVEYIPKGMKVF 290
Db 338 VLTELDSSILPLYQCVAKELR-----LHPPTPLM.PHRANANVYIKGYDIPKGSVH 390
QY 291 PIFTAVHLDPSLHNEPFEFNPMT-----KTAFGGAVRCPGEGELKQIAF 339
Db 391 VNVAVARDPAVWDRPLEFRERSEDDVDMKGDHYRLPLPGAGRRVCPGALGGINLVTIS 450

OY 340 ELHHLVSYRWK---IKSDMPIA 360
DB 451 MGHLLHFWYSPKGVKPEIDMS 475

RESULT 11

US-09-380-420C-2

Sequence 2, Application US/09380420C

Patent No. 6300544

GENERAL INFORMATION:

APPLICANT: Baikier, Barbara

Bak, Soren

Kahn, Rachel

Moller, Birger

TITLE OF INVENTION: Cytochrome P450 Monooxygenases

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Syngenta Patent Dept.

STREET: 3054 Cornwallis Road

CITY: RTP

STATE: NC

COUNTRY: USA

ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/380,420C

FILING DATE: 12-NO. 6300544-1999

CLASSIFICATION: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241

REFERENCE/DOCKET NUMBER: S-21251A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 531 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-380-420C-2

Query Match 8.5%; Score 172; DB 4; Length 531;
Best Local Similarity 22.3%; Pred. No. 2.3e-09;

Matches 87; Conservative 65; Mismatches 142; Indels 96; Gaps 17;

OY 50 DOELNMFILQ-----NEGKLTSDYPRAMHD-ILGYSLLATGEIH-----RKLK 94
DB 172 EQEMRLVADLDRAAASASIVLNDHVALTDCITG---TVAFGNTIASKQFAKKEKFQ 227
OY 95 NVITISFINL--TKSRPDLHCAENLSISLKSWMKNCREVEFHKEVKMFTLSVYNQLLSI 152
DB 228 HVLDDAMDMMAFSASDEFPNNAAGRLADRLSGFLARREIFN-ELDVFEKVIDQHM--- 283
OY 153 KPEDPARLYVLDPLSYMKGFISLPILPGTYTNAI-----KYRSNNITONAI 203
DB 284 ---DAR-----FVPDNGDVLVVLINLCKEHGTLRFTRD-HVKAIV 322
OY 204 ED-----MNNAIREDPLDSTISNDEDEHAIKAKGDGE-LLMEDYQ 246
DB 333 LDTFGALDTSVTLMMASMLMRKPYL-----RKAQLEVRVAAYDDKPRVNSEDA 375
OY 247 KMEFTQCYISEALRC-GNIVKTVHKKATHDIKKFEYVI PKGKVPYRPTVAHLDPSLHEN 305
DB 376 KIPYAKMVKETLRLHAPATLLVPREIMRDTTIGYDVPAITRVFVNMAAGRDASMPA 435
OY 306 PFEFNPMT-----KTFAFGGVYVPCGELGKLIQIAFFLHLVLSYRWK I-- 352

DB 436 PBEFNDRFVSGDVYXGSHFELIFGACRRCPLTGTETVTTLANLTCYDMALPG 495
OY 353 ---KSDMPIAH-YYEFRKMLLEIPTKF 379
DB 496 AMKPEDVSMEETGALTFRHKTPLVVVPKX 525

RESULT 12

US-08-532-065B-2

Sequence 2, Application US/08532065B

Patent No. 5753507

GENERAL INFORMATION:

APPLICANT: Ohla, Daisaku

Mizutani, Masaharu

TITLE OF INVENTION: Plant Geraniol/Nerol 10-Hydroxylase and

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 5753507artis Corporation

STREET: 59 Route 10

CITY: East Hanover

STATE: NJ

COUNTRY: USA

ZIP: 07936

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/532,065B

FILING DATE: 22-SEP-1995

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 495 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-532-065B-2

Query Match 8.4%; Score 170.5; DB 1; Length 495;
Best Local Similarity 19.7%; Pred. No. 2.9e-09;

Matches 88; Conservative 65; Mismatches 180; Indels 113; Gaps 14;

OY 4 PRIGETISFER-PHRSDSIGTFLOQRRVSRGKVRKSNICGKAVASCDQELNMFILQNEG 62
DB 46 PIGINHVLGNHPRHS-----FAELSKTYGPMVSLGSLNTVVAISPEAARVRLTHD 99
OY 63 KLFTSDYP-----KAMHDILGKYSU-----LATGEIHKKLK 94
DB 100 QILSARSTNVRISINODASLVLPSSSARMLRLRLSVYQLLSPORTEFTKALRNKV 159
OY 95 NVITISFINLTKSRPDLHCAENLSISLKSWMKNCREVEFHKEVKMFTLSVYNQLLSI 152
DB 160 KELVSFISESDRESDVISRVAFTTLNLSN---ILFSVDGSYNAKKSINGVQDTVI 216
OY 153 KPEDPARLYVLDPLSYMKGFISLPILPGTYTNAI-----QNAITE 204
DB 217 SYMDAAGTPDAANTFPLR-PLDL-----QGNVTEKVCYTERLVVRFQGTDAKIAKS 269
OY 205 DMNNA--IREDFLDSI-----ISNDEEH----- 227
DB 270 SONNPKDVSXNDYFDNLDLKYGDSELSISDIIEHLDDMTACTDTSSITLWMTLLK 329

```

OY 228 -----AATRAKGDCELLMEDYOKMEFTQCYSALRCGNIVK-TYHRKATHDIX 277
DB 330 NPKTAKQAQAEIDCVYIGONGIVEESDISKLPYLAQVAKETFRLLHPVLLIPRKAESDAE 389
OY 278 FKEYVYIPKGMKVFPIFTAVHLDPSSLHENPFEPNPMRW------KTTAFGGGVAV 326
DB 390 ILGFVAVLKDQVLVAVVMAIGRDPSPWMDNPSQEPERPERFLGKMDVGRDDELIPFGARRI 449
OY 327 CPGGELGKLQIAFPLHLHLVLSYRWKXI 352
DB 450 CPGMPLAMKTVSLMLASLISYSPDWKL 475

RESULT 13
US-08-313-075A-38
; Sequence 38, Application US/08313075A
; Patent No. 5639870
; GENERAL INFORMATION:
; APPLICANT: Holton, Timothy A.
; APPLICANT: Cornish, Edwin C.
; APPLICANT: Tanaka, Yoshihazu
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID
; TITLE OF INVENTION: PATHWAY ENZYMES AND USES THEREFOR
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,075A
FILING DATE: 30-NOV-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PL 1538/92
FILING DATE: 27-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PL 6698/93
FILING DATE: 07-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PCT/AU93/00127
FILING DATE: 25-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9433
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-313-075A-38

Query Match 8.4%; Score 169.5; DB 1; Length 506;
Best local Similarity 21.3%; Pred. No. 3.9e-09;
Matches 100; Conservative 64; Mismatches 154; Indels 151; Gaps 24;
OY 2 GWPPIGE-TISFRRPHRSDSIGTFLQORVSRKGV--FKSNICGGAIVVSCDDELNFIIL 58
DB 39 GWPVIGALPILGAMPVHS-----LAKMARKYGAIMYLKVGTC-GMAVASTPDAKAF-L 90

```

```

OY 59 QNEGKLFSTDYPK--AMHDILGKYSLLA-----TGEIHRKLKNVLIISFINLTKSKPDFLH 112
DB 91 KTLIDNFSRRPPNAGATILANADMAVFAHGPRAKILKLSN-----LH 135
OY 113 CAENLSISILSKWKNCREVER-H-----KEVKFTLSVAVNOLLSTKP 154
DB 136 M---LGKALEWMAVAVRANELGHMLKSMDSMREGQRYVVAEMLIFFAANNIGOMLSK- 191
OY 155 EDPALVY-----LQDF-----LSYMKGISLPPIPGTGTN--AIKVSNNRIHQ-- 199
DB 192 ---RVFVDKGYEVNEFKDMVVELMTIAGYFNIGDFIPLAMMDLGITEKMKR-LHKKF 246
OY 200 NAIIDMNNAIR-----EEDFLDSIISNEDEHAA-----IRA----- 232
DB 247 DALLIKMDEHKATYERKGRKDPIDVYMENDNGSEBELSTYNKALLNLFTAGTDTS 306
OY 233 -----KKGDEL-----LMEDYOKMEFTQCYSALRCG--- 262
DB 307 SSAIEMALAEKMKNPAILKKAQAEQDQYIGRRNRILLESIDIPMLPYLRAICKETFRKHST 366
OY 263 --NIYKTVARKATHDIKFEYVYIPKGMKVFPIFTAVHLDPSSLHENPFEPNPMRW- 315
DB 367 PLNLPRIISNEPCTVD---GTYIPKNTRLSVIWAIGRDPQVWENPLFEPNPERFLSGANS 422
OY 316 -----KTTAFGGGVAVCPGGELGKLQIAFPLHLHLVLSYRWKIKSD 355
DB 423 KIDPRGNDELIPFGARRICAGTIRMGIVMEYILGTLVHSPDWKLPSB 471

RESULT 14
US-08-606-505B-65
; Sequence 65, Application US/08606505B
; Patent No. 6114601
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yasuhiro
; APPLICANT: KIKOKAWA, Shigeto
; APPLICANT: SHIMADA, Yukihisa
; APPLICANT: OHBAYASHI, Masaya
; APPLICANT: SHIMADA, Ritsuko
; APPLICANT: OKINAKA, Yasushi
; TITLE OF INVENTION: NOVEL PLANT GENES
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FITZPATRICK, CELIA, HARPER & SCINTO
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112-3801

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 KB storage.
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/606,505B
FILING DATE: 23-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Perry, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200
INFORMATION FOR SEQ ID NO: 65 :
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

```


MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Petunia hybrida
STRAIN: Falcon Blue
FEATURE:
NAME/KEY: CDS
LOCATION: 116 to 1633
IDENTIFICATION METHOD: by experiment
US-08-606-505B-65

Query Match 8.4%; Score 169.5; DB 3; Length 506;
Best Local Similarity 21.3%; Pred. No. 3.9e-09;
Matches 100; Conservative 64; Mismatches 154; Indels 151; Gaps 24;

QY 2 GMPFGE-TISFFKPHSDSIGTLOQVRSYGV--FKSNIGGKAVVSCDDELNFI 58
DB 39 GMPVIGALPLLGAMPVHS-----LAKMAKKYGALIMLKVGTC-GMAVASTPPAKAF-L 90
QY 59 QNEGKLFSDYPR--AMHDLGKYSLLA---TGEIHRKLKAVIISFINLTGSKPDFLH 112
DB 91 KTLIDINSNRPNGATHLAVNADQVFAHYGPRMKLLRKLSTN-----LH 135
QY 113 CAENLSISILSKMKNCREVEF-H-----KEYKMTLSVNVQLSTIRP 154
DB 136 M---LGKALENMANVANLGHMLKMSDMSREGORVVAEMLTFRAMNMIGQVWLK- 191
QY 155 EDPAFLYV-----LQDF-----LSYMKGFISLPPLPGTGYN--AIKVRNRIHQ-- 199
DB 192 ---RVFVDKGVENRKKDMVVELMTAGYFNIGDFIPCLAMNDLOGIEKRMKR-LHKKF 246
QY 200 NAIEDNNNAIR-----EEDFLDSISNDEEHA-----IA----- 232
DB 247 DALLTKMFDEHKATYTERKCKRPFLDYVMENGNSGERSLSTTNIALLLNFTAGTDS 306
QY 233 -----KKGDEL-----LMWEDYQKMEFTQCVISEALRCG--- 262
DB 307 SSAIEMALAEKMKNPAILKKAQAEQVIGRNRRLLESIDIPNLPYLAICKETFRKHST 366
QY 263 --NIVTVHRRKATHDIKFEYVIRPKMKVFPFTAVHLDPSLHENFEENPRKMT----- 315
DB 367 PLNLPRISNEPCVD---GYITPKNTRLSVINMAIGRPOVWENPLENPRFLSGRNS 422
QY 316 -----KTTAFGGGVNCPGEGELGKQIAFFLHNLVLSYRWIKSD 355
DB 423 KIDPRGNDFELIPFGAGRICAGTRMGIVMVEYIIGTLVHSFDMKLPS 471

RESULT 15
US-09-616-990-65
Sequence 65, Application US/09616990
Patent No. 6232109

GENERAL INFORMATION:
APPLICANT: KIKUCHI, Yasuhiro

KIKOKAWA, Shigeto
SHIMADA, Yukihisa
OHBAAYASHI, Masaya
SHIMADA, Ritsuko
OKINAKA, Yasushi

TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:

ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza

CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
COMPUTER: IBM PS/4
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/616,990
FILING DATE: 14-Jul-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Peiray, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Petunia hybrida
STRAIN: Falcon Blue
FEATURE:
NAME/KEY: CDS
LOCATION: 116 to 1633
IDENTIFICATION METHOD: by experiment
SEQUENCE DESCRIPTION: SEQ ID NO: 65
US-09-616-990-65

Query Match 8.4%; Score 169.5; DB 4; Length 506;
Best Local Similarity 21.3%; Pred. No. 3.9e-09;
Matches 100; Conservative 64; Mismatches 154; Indels 151; Gaps 24;

QY 2 GMPFGE-TISFFKPHSDSIGTLOQVRSYGV--FKSNIGGKAVVSCDDELNFI 58
DB 39 GMPVIGALPLLGAMPVHS-----LAKMAKKYGALIMLKVGTC-GMAVASTPPAKAF-L 90
QY 59 QNEGKLFSDYPR--AMHDLGKYSLLA---TGEIHRKLKAVIISFINLTGSKPDFLH 112
DB 91 KTLIDINSNRPNGATHLAVNADQVFAHYGPRMKLLRKLSTN-----LH 135
QY 113 CAENLSISILSKMKNCREVEF-H-----KEYKMTLSVNVQLSTIRP 154
DB 136 M---LGKALENMANVANLGHMLKMSDMSREGORVVAEMLTFRAMNMIGQVWLK- 191
QY 155 EDPAFLYV-----LQDF-----LSYMKGFISLPPLPGTGYN--AIKVRNRIHQ-- 199
DB 192 ---RVFVDKGVENRKKDMVVELMTAGYFNIGDFIPCLAMNDLOGIEKRMKR-LHKKF 246
QY 200 NAIEDNNNAIR-----EEDFLDSISNDEEHA-----IA----- 232
DB 247 DALLTKMFDEHKATYTERKCKRPFLDYVMENGNSGERSLSTTNIALLLNFTAGTDS 306
QY 233 -----KKGDEL-----LMWEDYQKMEFTQCVISEALRCG--- 262
DB 307 SSAIEMALAEKMKNPAILKKAQAEQVIGRNRRLLESIDIPNLPYLAICKETFRKHST 366
QY 263 --NIVTVHRRKATHDIKFEYVIRPKMKVFPFTAVHLDPSLHENFEENPRKMT----- 315
DB 367 PLNLPRISNEPCVD---GYITPKNTRLSVINMAIGRPOVWENPLENPRFLSGRNS 422
QY 316 -----KTTAFGGGVNCPGEGELGKQIAFFLHNLVLSYRWIKSD 355
DB 423 KIDPRGNDFELIPFGAGRICAGTRMGIVMVEYIIGTLVHSFDMKLPS 471

Search completed: July 29, 2002, 13:58:24
Job time: 51 sec

ALIGNMENTS

Tue Jul 30 08:40:07 2002

us-09.

```
RESULT 1
ID Q9LY89 PRELIMINARY; PRT; 382 AA.
AC Q9LY89;
DT 01-OCT-2000 (TREMBLrel. 15, created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)
GN HYPOTHECAL 43.9 KDA PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Euarystea viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsid.
OK NCBI_taxid=3702;
RN (1)
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Hewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RA ED Arabidopsis sequencing project:
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL, AL163817; CAB8779.1; -.
DR InterPro: IPR001128; CYP_P450.
DR Pfam: PF00067; P450; 1.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN 1.
KW Heme; Hypothetical protein; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 382 AA; 43889 MW; 1BC565AB73B4E30 CRC64;
```

```
QY 61 ESKLFTSYPRAMHDIIGKYSLLATGELHRRKLKNVTSFINITKSPDFLCAENLSTIS 120
DB 61 ESKLFTSDYPRAMHDIIGKYSLLATGELHRRKLKNVTSFINITKSPDFLCAENLSTIS 120
QY 121 ILKSNKNCREVEFEKFKMFTLSVMNQLLSTPEEDPARLYLODFLSYMGFISLPPL 180
DB 121 ILKSNKNCREVEFEKFKMFTLSVMNQLLSTPEEDPARLYLODFLSYMGFISLPPL 180
QY 181 PGTGYTNAIKVRSNRNIHQAIIEDMNNAIREDPFLDSTISNDEDEHAAIRAKKDGCEL 240
DB 181 PGTGYTNAIKVRSNRNIHQAIIEDMNNAIREDPFLDSTISNDEDEHAAIRAKKDGCEL 240
QY 241 NMEDYOKMEFTQCVISEALRCGNIVYTKVTRKATHDIKREYVTPGKRVFPITFAVHDP 300
DB 241 NMEDYOKMEFTQCVISEALRCGNIVYTKVTRKATHDIKREYVTPGKRVFPITFAVHDP 300
QY 301 SLHENPEERPNRMTKTAFGGGVNCPGEGELGLQIAFLHHLVLSYRKIKSDMPIA 360
DB 301 SLHENPEERPNRMTKTAFGGGVNCPGEGELGLQIAFLHHLVLSYRKIKSDMPIA 360
QY 361 HPYVEFRKMLLEIPTKFLD 382
DB 361 HPYVEFRKMLLEIPTKFLD 382
```

```
Query Match          99.68; Score 2018; DB 10; Length 382;
Best Local Similarity 99.58; Pred. No. 1,le-153;
Matches 380; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 MCMPIGETISFFKPHRSDISGTFLQORVSRYGKFKSNIGCGKAVVSCDDELNMFTLION 60
DB 1 MCMPIGETISFFKPHRSDISGTFLQORVSRYGKFKSNIGCGKAVVSCDDELNMFTLION 60
```

- Support yield this is it aligned?

- Support function based on spec

- % 500 is of function

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 13:57:33 : Search time 28.42 seconds
(without alignments)
2325.266 Million cell updates/sec

Title: US-09-995-917A-1
Perfect score: 2027
Sequence: 1 MGPPIGETISFFKPHRSDS.....YVEFKRGMLLEIETPFLLED 382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPTRMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2018	99.6	382	10	Q9LY89
2	822	40.6	513	10	Q64989
3	821	40.5	513	10	Q9SC09
4	642.5	31.7	474	10	Q9LKH7
5	599.5	29.6	512	10	Q9FX29
6	592.5	29.2	491	10	Q94IA6
7	586	28.9	490	10	Q94IM5
8	573	28.3	464	10	Q9LIC5
9	572	28.2	465	10	Q94OV4
10	572	28.2	478	10	Q9LW73
11	569	28.1	465	10	Q9LH81
12	546	26.9	465	10	Q9FMA5
13	513.5	25.3	467	10	Q949P1
14	513	25.3	463	10	Q9FT38
15	505.5	24.9	463	10	Q9FH76
16	504	24.9	457	10	Q65624

17	487.5	24.1	485	10	Q9SJH2	Q9SJH2 arabidopsis
18	486	24.0	443	10	Q9LJK2	Q9LJK2 arabidopsis
19	484	23.9	482	10	Q81077	Q81077 arabidopsis
20	468.5	23.1	735	10	Q9LG17	Q9LG17 arabidopsis
21	467	23.0	486	10	Q9LIE9	Q9LIE9 arabidopsis
22	460	22.7	455	10	Q9LKH8	Q9LKH8 arabidopsis
23	431.5	21.3	477	10	Q9LVY7	Q9LVY7 arabidopsis
24	427	21.1	489	10	Q9ZV72	Q9ZV72 arabidopsis
25	427	21.1	489	10	Q9CSY2	Q9CSY2 arabidopsis
26	423	20.9	490	10	Q9CSY3	Q9CSY3 arabidopsis
27	422.5	20.8	464	10	Q04949	Q04949 arabidopsis
28	422.5	20.8	474	10	Q9SHY7	Q9SHY7 arabidopsis
29	417.5	20.6	496	10	Q9FQY4	Q9FQY4 cucurbita m
30	416	20.5	487	10	Q23384	Q23384 arabidopsis
31	413	20.4	497	10	Q9AXM6	Q9AXM6 taxus cuspi
32	412.5	20.4	460	10	Q9SVX2	Q9SVX2 arabidopsis
33	408.5	20.2	504	10	Q9SNC3	Q9SNC3 oryza sativ
34	376	18.5	499	10	Q9AXH9	Q9AXH9 hordeum vul
35	292.5	14.4	497	11	Q9R1F4	Q9R1F4 mus musculu
36	283.5	14.0	224	10	Q9SDM6	Q9SDM6 helianthus
37	280.5	13.8	375	10	Q9LW32	Q9LW32 arabidopsis
38	280	13.8	492	13	Q93323	Q93323 xenopus lae
39	278.5	13.7	512	4	Q9NR63	Q9NR63 homo sapien
40	276	13.6	525	4	Q9NP41	Q9NP41 homo sapien
41	275	13.6	492	13	Q9PUB4	Q9PUB4 gallus gall
42	224.5	11.1	494	11	Q9QZ50	Q9QZ50 mesocricetu
43	223.5	11.0	494	11	Q91XG2	Q91XG2 mus musculu
44	217.5	10.7	494	11	Q91X75	Q91X75 mus musculu
45	215.5	10.6	349	10	Q9FMR7	Q9FMR7 arabidopsis

ALIGNMENTS

RESULT 1
Q9LY89 PRELIMINARY: PRT: 382 AA.
ID Q9LY89;
AC Q9LY89;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DR 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOHETICAL 43.9 KDA PROTEIN.
GN F18022_190.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL: AL163817; CAB87779.1; -;
DR InterPro: IPR001128; Cyt_P450.
DR Pfam: PF00067; P450.1.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN.1.
KW Heme: Hypothetical protein: Monooxygenase: Oxidoreductase.
SQ SEQUENCE 382 AA; 43889 MW; 1B5685AB73BAE30 CRC64;

Query Match 99.6%; Score 2018; DB 10; Length 382;
Best Local Similarity 99.5%; Pred. No. 1,1e-153;
Matches 380; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPIGETISFFKPHRSDSITFLQQRVSRGKVKFSGNCGKAVVSDQDLNMFILN 60
DB 1 MGPPIGETISFFKPHRSDSITFLQQRVSRGKVKFSGNCGKAVVSDQDLNMFILN 60

```

QY 61 ECKLTSDYKAMHDLGKYSLLATGELTHRKLVNIIISFINLTKSPDLHCAMLSIS 120
DB 61 ECKLTSDYKAMHDLGKYSLLATGELTHRKLVNIIISFINLTKSPDLHCAMLSIS 120
QY 121 ILKSNKNCREVEFHKEVKFTLSVMVNOQLLSIKPEDPARLYVLODFLSYMKGFISLPIL 180
DB 121 ILKSNKNCREVEFHKEVKFTLSVMVNOQLLSIKPEDPARLYVLODFLSYMKGFISLPIL 180
QY 121 ILKSNKNCREVEFHKEVKFTLSVMVNOQLLSIKPEDPARLYVLODFLSYMKGFISLPIL 180
DB 121 ILKSNKNCREVEFHKEVKFTLSVMVNOQLLSIKPEDPARLYVLODFLSYMKGFISLPIL 180
QY 181 PGTGTNAIKVSNRNHONAIIEDMNNAIREDFDLSITSNEDEHAAIRAKKGDELL 240
DB 181 PGTGTNAIKVSNRNHONAIIEDMNNAIREDFDLSITSNEDEHAAIRAKKGDELL 240
QY 241 NMEDYQKMEFTQCVISEALRCGNIVTKRKATHDIKFKEYIIPKGMKVPPIFTAVHLDP 300
DB 241 NMEDYQKMEFTQCVISEALRCGNIVTKRKATHDIKFKEYIIPKGMKVPPIFTAVHLDP 300
QY 301 SLHNEPFEFNPWRMTKTAFAGGGVRCVCPGELGKLQIAFELHHLVLSYRKIKSDMPA 360
DB 301 SLHNEPFEFNPWRMTKTAFAGGGVRCVCPGELGKLQIAFELHHLVLSYRKIKSDMPA 360
QY 361 HPYVEFKRGMLLEIETKFLD 382
DB 361 HPYVEFKRGMLLEIETKFLD 382

```

RESULT 2

```

ID 064989 PRELIMINARY: PRT: 513 AA.
AC 064989;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE STEROID 22-ALPHA-HYDROXYLASE.
GN DWF4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RM SEQUENCE FROM N.A.
RC STRAIN=MS-2;
RX MEDLINE=98158690; PubMed=9490746;
RA Choe S., Dilkes B.P., Fujioka S., Takatsuto S., Sakurai A.,
RA Feldmann K.A.;
RT "The DWF4 gene of Arabidopsis encodes a cytochrome P450 that mediates
RT multiple 22alpha-hydroxylation steps in brassinosteroid
RT biosynthesis";
RL Plant Cell 10:231-243(1998).
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF044216; AAC05093.1; -
DR InterPro; IPR001128; Cyt_P450.
DR Pfam; PF00067; P450.1.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 513 AA: 58867 MW; B1639BDD9A4DA6F3 CRC64;

```

Query Match 40.6%; Score 822; DB 10; Length 513;
 Best Local Similarity 34.4%; Pred. No. 1,le-57;
 Matches 160; Conservative 87; Mismatches 126; Indels 92; Gaps 4;

```

QY 2 GMPFGETISFEKPHRSISIGTFLOQVRSRYKGVKSNICGKAVVSCQOELNMFLONE 61
DB 45 GMPFGETISFEKPHRSISIGTFLOQVRSRYKGVKSNICGKAVVSCQOELNMFLONE 61
QY 62 GRLTSDYKAMHDLGKYSLLATGELTHRKLVNIIISFINLTKSPDLHCAMLSIS 121
DB 105 GRLTSDYKAMHDLGKYSLLATGELTHRKLVNIIISFINLTKSPDLHCAMLSIS 121
QY 122 LKSNKNCREVEFHKEVKFTLSVMVNOQLLSIKPEDPARLYVLODFLSYMKGFISLPIL 181
DB 122 LKSNKNCREVEFHKEVKFTLSVMVNOQLLSIKPEDPARLYVLODFLSYMKGFISLPIL 181

```

```

DB 165 LDSQOONISFSAQDEAKFTFNLMAKHMSMDGEEETEOLKREYVFMKGVSAADNLDP 224
QY 182 GTCYTAIKVSNRNHONAIIEDMNNAIREDFDLSITSNEDEHAAIRAKKGDELL 240
DB 225 GTCYTAIKVSNRNHONAIIEDMNNAIREDFDLSITSNEDEHAAIRAKKGDELL 240
QY 225 GTCYTAIKVSNRNHONAIIEDMNNAIREDFDLSITSNEDEHAAIRAKKGDELL 240
DB 225 GTCYTAIKVSNRNHONAIIEDMNNAIREDFDLSITSNEDEHAAIRAKKGDELL 240
QY 285 DILGWYKHSNLSLTDQILDLISLFAGHETSSVATALAIFFLOQPKAVEELREHLEI 344
DB 285 DILGWYKHSNLSLTDQILDLISLFAGHETSSVATALAIFFLOQPKAVEELREHLEI 344
QY 231 -RAKKGDELL-LNEDYQKMEFTQCVISEALRCGNIVTKRKATHDIKFKEYIIPKGMK 288
DB 345 ARAKKELGSESLNWDYKMDFTQCVINETLRIGNVVRFLHRKALDVRKGYDIPSGMK 404
QY 289 VPFIFTAVHLDPSTLHNEPFEFNPWRMTKTAFAGGGVRCVCPGELGKLQIAFELHHLVLSYRKIKSDMPA 360
DB 405 VPFIFTAVHLDPSTLHNEPFEFNPWRMTKTAFAGGGVRCVCPGELGKLQIAFELHHLVLSYRKIKSDMPA 360
QY 330 GELGKLQIAFELHHLVLSYRKIKSDMPAHPYVEFKRGMLLEI 374
DB 465 SELAKLEMAVFIHHLVLFKFMELAEDDQPPAFPFVDFPNGLPIRV 509

```

RESULT 3

```

ID 09SC09 PRELIMINARY: PRT: 513 AA.
AC 09SC09;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE STEROID 22-ALPHA-HYDROXYLASE (DWF4) (AT3G50660/73A5_40).
GN 73A5_40.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RM SEQUENCE FROM N.A.
RA Bloeker H., Mewes H.W., Lemcke K., Mayer K.F.X., Quetier F.,
RA Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RM SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RM SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Tortum M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AL132979; CAB62435.1; -
DR EMBL; AF412114; AAL06567.1; -
DR InterPro; IPR001128; Cyt_P450.
DR Pfam; PF00067; P450.1.
DR PRINTS; PR00365; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 513 AA: 58867 MW; B1639BDD9A5D7C93 CRC64;

```

Query Match 40.5%; Score 821; DB 10; Length 513;
 Best Local Similarity 34.4%; Pred. No. 1,3e-57;
 Matches 160; Conservative 87; Mismatches 126; Indels 92; Gaps 4;

```

OY 2 GMPFGETISFFKPHRSISGTFLOQSVRSYKVEKSNICGKAVSCDOELNMFLLONE 61
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 45 GMPFGETIGYTLKPYATATLGDPMOONHVSKEYGIRYSNLFGEPTIYSADAGLNFILONE 104
OY 62 GKLFTSDYKAMHDLGKSLLATGEIRKLKNVISTINTLTKSPDPLHCAENLSISI 121
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 105 GKLFTSCYPRISIGLIGKMSMLVAGDMHDMRSISLNLSHARLRTLLKDKVERHTLTV 164
OY 122 LKSMKNCREVEFHKEKMFLLSVNQLISIKPEDPARLYVQDFLSYWKGFISPIPLP 181
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 165 LDMQONSIFSAQDEKKTFTNLMAKHIMSDRGEETQOLKEVYTFPKGYVASAPLNP 224
OY 182 GCGYNAIKVRSNRNINHQAITEEDNNAIREDPLDISISNED----- 224
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 225 GYAHKALOSRATILKETERKMEERKLDIKEDEDEEEVKTEDAEKMSKDHVRKQRTDD 284
OY 225 -----BEHAI 230
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 265 DLGNVLKHSNSTEQIDLLISLFPAGHETSVAIALAIFLQACPKRAVELREHLEI 344
OY 221 -RAKKGDEL-LNMDYOKMEFTQCVISALRCGNIVTKVRKATHDIEKFEYVIPKWK 288
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 345 ARAKKELESESLNMDYKMDFTQCVINETLRLGNVVRFLRKALKDVKYKGYDIPSGWK 404
OY 289 VPFITAVHLDPSLHENPPEFPMRWTKT-----AFGGGVAVCPG 329
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 405 VLPVISAHLDSRYDQPLNFPWRMOONNGASSSGSGSFTMGNNYMPFGGPRLAG 464
OY 330 GELGKIQIAFLHLVLVSYRMKIKSDMPIAHRYVEFKGMLEI 374
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 465 SELALBMAVFIHLVLKFNMLADDKFAFPVDFPGLDIRV 509

```

Query Match 31.7% Score 642.5; DB 10; Length 474;
 Best Local Similarity 32.7%; Pred. No. 2.4e-43;
 Matches 140; Conservative 71; Mismatches 140; Indels 77; Gaps 6;

```

OY 2 GMPFGETISFFKPHRSISGTFLOQSVRSYKVEKSNICGKAVSCDOELNMFLLONE 61
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 37 GMPFGETIGYTLKPYATATLGDPMOONHVSKEYGIRYSNLFGEPTIYSADAGLNFILONE 96
OY 62 GKLFTSDYKAMHDLGKSLLATGEIRKLKNVISTINTLTKSPDPLHCAENLSISI 121

```

```

Db 97 GKLDCSPGSISSLNLGKHSLLMGALHKRHSLSLMSRANSISIKDHLHIDNLIGLN 156
OY 122 LKSMKNCREVEFHKEKMFLLSVNQLISIKPEDPARLYVQDFLSYWKGFISPIPLP 181
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 157 LDTWSD--RVTLMDQAKKITFELTVKQLSMFPDDEWTE-SLRKEVVLVIEGFFTLPLP 213
OY 182 GCGYNAIKVRSNRNINHQAITEEDNNAIREDPLDISISNED-----EFDLSISNED----- 224
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 214 SITTYRAIKARK-----VAEALTLVQRREYNOCKEKKSMPLALASGHSFSD 266
OY 225 -----BEHAI 230
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 267 QIVDFLLALVAGYETSTIMTLAVKFLTERPLALQKLKEHDQIRASDPGAPLEMTDY 326
OY 246 QKMEFTQCVISALRCGNIVTKVRKATHDIEKFEYVIPKWKVPITVAVHLDPSLHEN 305
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 327 KSMVETQHVNETLKVANIIGIFRRATTDIDIKCYTIPKWKVFASPRAVHLNDEYYKD 386
OY 306 PPEFPMRWTKT-----TAFGGGVAVCPGSELGKIQIAFLHLVLVSYRMKIKS 354
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 367 ARTEFPMRWNOSSSRAAMPANVYTFPGGPRLCFVELARVYSLVFLHRTYRISWPAE 446
OY 355 DEMPAAHP 362
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 447 EDKLVFFP 454

```

Query Match 29.6% Score 599.5; DB 10; Length 512;
 Best Local Similarity 30.6%; Pred. NO. 7.4e-40;
 Matches 144; Conservative 74; Mismatches 139; Indels 113; Gaps 15;

```

OY 2 GMPFGETISFFKPHRSISGTFLOQSVRSYKVEKSNICGKAVSCDOELNMFLLONE 61
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 49 GMPFGETIGYTLKPYATATLGDPMOONHVSKEYGIRYSNLFGEPTIYSADAGLNFILONE 96
OY 37 KSNICGKAVSCDOELNMFLLONEGKSLLATGEIRKLKNVISTINTLTKSPDPLHCAENLSISI 121
    ||| ||| : : : ||| ||| : : : ||| : : : ||| : : : ||| : : : |||
Db 109 SCSEFGKMAVVASADPDRNFRIMQNEKGLFQSSYKSPSRDLVKGKGVITVHNDQQRRLHSI 168
OY 97 IISINTLTKSPDPLHCAENLSISILKSMKN-----CREVEFHKEKMFLLSVN 147

```

Db 169 ASSSMHHDLKTHFEVIVPMLOITSNKDEEVLLDIDICKVAIH-----LWVN 219
Oy 148 QLLSIKPEDPARLYVLODFLSYMKGFISLPILPGTGTNAIKVRS-----NRNIH--- 198
Db 220 QLLGYSSESEVD-EMSQLFSDFVDGCLSVIPDLPGFTYKAKAKREILRNKITEKRL 278
Oy 199 QNALIED-----MNNNAIEE-----DFLDSIT--SNED----- 224
Db 279 QNKASDFTANGVGLRLLEEESLPNESMADFTINLFAENETTSKTMLEPAVYFLTHCPKA 338
Oy 225 -----EEHAIRAKKGDGKEDLNMEDYQKMEFQCVISEALRCGNIVKTVHRKATHDIKFK 279
Db 339 MTQLLEFHDRLA-----GGMLTWQDYKTMDFTOCVIDEFLRLGGIAIWLMEKAEKVSTQ 393
Oy 280 EYVIPKGMKVFPLFTVAHLDPSLHENPEFENPMRWTF-----KTT-----AEGGYVR 325
Db 394 DYVIPKGCVFVPLSAVHLDESYYKESLSEFPMRWMLDPETQOKRMRTSPFCFGGTR 453
Oy 326 VEPGGEGLKQIAFLFHLVLSIRW-KIKSDEPRLAHPIVEERKMLLEI 374
Db 454 FCPGAEIARLQIALFLHYFTTYKMTQLEKDRISF-FPSARLVNGFKIQL 502

RESULT 6
O94IAG PRELIMINARY; PRT; 491 AA.
AC 094IAG;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE CYP90D.
GN CYP90D.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Shimada Y.;
RT *P450 gene repressed by brassinosteroid.*;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB066286; BAB62109.1; - P450/GenBank/DBJ databases.
SQ SEQUENCE 491 AA; 56153 MW; 02FB908A91995A40 CRC64;

Query Match 29.2%; Score 592.5; DB 10; Length 491;
Best local Similarity 31.4%; Pred. No. 2.6e-39;
Matches 137; Conservative 75; Mismatches 160; Indels 65; Gaps 8;
Oy 1 GMPFIEGTISFEKPHRSDSIGTFLOQVRYSRYKVKFSNIGCGKAVYSCDQELNMTLON 60
Db 56 LGPVIETIEFYSASYSRPSPEMDKRLRMVGRFVSHLFTGATITVSDAEVNRVLOS 115
Oy 61 EGGLETSVPKAMHDLGKYSLLATGELIRKLNVIISFINLTKSKPDLHCAENLIS 120
Db 116 DSTAFVPTFYKYRELKSGSIILLINGSILRRHGLVGSFLSKPLKAQIVROMHKLSE 175
Oy 121 ILKSMKNCREVEFHKEVKMFTLSVMVNOQLSI-KPEDPARLYVLODFLSYMKGFISLPILP 179
Db 176 SMDLMSDQVYLLQDYSKVAFAFKVALKALISVEKGEDEEL--KREPFNISGLMSIPIN 233
Oy 180 LPETGTINAL-----KVRNRRNINHONAIT-----EDMNNAI 210
Db 234 FPGTOLHRSLOAKKMMKQYERITTEGKIRTKKREEDVDIAKDVVDLLKDSSEHLTHNL 293
Oy 211 REDPFLDSIISNED-----EEHAIRA-KKGDGELLNMEY 245
Db 294 IANNMIDMILPGHDSVAVLTLAVKFLSDSPALNLTENMKLSIKELTGEPLVWYNDY 353
Oy 246 QKMEFTQCVISEALRCGNIVKTVHRKATHDIKREYVIPKGMKVFPLFTVAHLDPSLHEN 305

Db 354 LSLPFTQKIYITTLKMGVNIITGMKAKMDVIEIKGVIPKGCFLAIRSVLIDLTYES 413
Oy 306 PEEFNPMEVN-----TKTTAFGGVAVRCPGGEGLKQIAFLFHLVLSYRKIKSDEMP 358
Db 414 PYKFNPMWRQERDMNTSSPSFGGQRLCPGLDLARLETSTVSLHLVIRFRW-IAEDPTI 472
Oy 359 IAHPIVEERKMLLEIE 375
Db 473 INFPIVHMKNKLPIWK 489

RESULT 7
O94IWS PRELIMINARY; PRT; 490 AA.
AC 094IWS;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE CYTOCHROME P450-LIKE PROTEIN.
GN P0419B01.5.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae.
OX Euhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsunoto T., Yamamoto K.;
RT *Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC
clone: P0419B01.*;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003244; BAB56089.1; - 9EC2853BBAFB88F CRC64;
SQ SEQUENCE 490 AA; 54824 MW; 9EC2853BBAFB88F CRC64;

Query Match 28.9%; Score 586; DB 10; Length 490;
Best local Similarity 30.4%; Pred. No. 8.5e-39;
Matches 131; Conservative 87; Mismatches 153; Indels 60; Gaps 9;
Oy 2 GMPFIEGTISFEKPHRSDSIGTFLOQVRYSRYK-VKFSNIGCGKAVYSCDQELNMTLON 60
Db 55 GMPVVEETIEFYSASCASPRPEAFVDRRLHGSVAFRSHLFGSATVVTADAEVSFRVLOS 114
Oy 61 EGGLETSVPKAMHDLGKYSLLATGELIRKLNVIISFINLTKSKPDLHCAENLIS 120
Db 115 DARAFVPTFPRSLTELKSGSIILLINGALQRRYHGLVGAFFKSHLKSQLTADMRRRLSP 174
Oy 121 ILKSMKNCREVEFHKEVKMFTLSVMVNOQLSIKPEDPARLYVLODFLSYMKGFISLPILP 180
Db 175 ALSFPDSSILHYQHLAKSVFEILVRLGLGLEAGEMO-QLKQOPQEFIVGLMSIPIKL 233
Oy 181 PGGTGTNAIKVRS-----NRNIHQ-----NAIIEDMNNNAIEE---DF 215
Db 234 FPGTLYRSLOAKKMMARLQRLIRERARRAASPPRDALIVLIGSDSELDLISDMK 293
Oy 216 LDSIISNED-----EEHAIRA-KKGDGELLNMEYQKMEF 250
Db 294 IDLMTAFEDSVYLLITLAVKFLSECPDLAHQLEENIQOLKRRKTDGCELTQWIDYMSLSE 353
Oy 251 TQCVISEALRCGNIVKTVHRKATHDIKREYVIPKGMKVFPLFTVAHLDPSLHENPEFN 310
Db 354 TQHVITETLRLGNIGIMKAKAVDEVGHLIPKGCWCFVFRSVHLDLDTLYDEPYKFN 413
Oy 311 PMRWTKT-----TAFGGVAVRCPGGEGLKQIAFLFHLVLSYRKIKSDEMPIAHY 363
Db 414 PMRWKEDMSNGSFPPGGQRLCPGLDLARLEASTIFLHLVLSFRVVAEEDHT-VNFTPL 472
Oy 364 VEFKRGMLLEI 374
Db 473 VRLKRGPIRV 483

ID	Seq1	Seq2	Score	DB	Length
09LIC5	PRELIMINARY;	PRF:	464	AA.	
AC	09LIC5				
DT	01-OCT-2000 (TREMBLrel. 15, Created)				
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)				
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DE	CYTOCHROME P450-LIKE PROTEIN				
OS	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;				
OC	Eurosidia II; Brassicales; Brassicaceae; Arabidopsidae.				
OX	NCBI_TaxID:3702;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-COLUMBIA;				
RC	Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.,				
RL	submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.				
RN	(2)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-COLUMBIA;				
RX	MEDLINE-20363099; PubMed-10907853;				
RA	Nakamura Y.;				
RT	*Structural analysis of Arabidopsis thaliana chromosome 3. II.				
RT	Sequence features of the regions of 4,251,695 bp covered by ninety pl,				
RT	TAC and BAC clones.;				
RL	DNA Res. 7:217-221(2000).				
CC	-1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.				
DR	EMBL; AP001307; BAB01922.1;				
DR	Interpro: IPR001128; Cyt_P450.				
DR	Interpro: IPR000504; RM.				
DR	Pfam: PF00067; P450. 1.				
DR	PRINTS: PR00365; P450.				
DR	PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.				
DR	PROSITE: PS00030; RM_RNP_1; UNKNOWN_1.				
KW	Heme; Monooxygenase; Oxidoreductase.				
SQ	SEQUENCE 464 AA; 52864 MW; A65E094665E51908 CRC64;				
Query Match	28.3%;	Score 573;	DB 10;	Length 464;	
Best Local Similarity	31.6%;	Freq. No. 6.7e-38;			
Matches 130;	Conservative 70;	Mismatches 147;	Indels 64;	Gaps	
QY	1	MGMPRTGETISFFKPHRSDSICFQQRSRGKFKFSNCGKAVVSCDDELNFILON 60			
DB	56	LGMPIGTETIEVSSAYSIDRSEFSDKRRKMGVGRFKSHIFGTATIVSDAEVNAVLDS 115			
QY	61	EGLKLTSDYPRKAMHIDLCYSLLATGELHKKLVAVIISFNLTKSKDFELCAENLST 120			
DB	116	DSTAVPVPYPTVRLKMKSSITLLNGSLHRRHFLVSEFLKSPILKQIVDMHKLPLSE 175			
QY	121	ILKSKNCRREVFEHKEVMEFTLSVWNOLST-KREDPARLVLDPLSMGFSILPT 179			
DB	176	SMDLMSDEQPVLDQVSTVAFKVLAKALISYEKDEDEEL-KREFENFISGLWSPLPN 233			
QY	180	LPGTGYTNAI-----KYRSNRNIIQNAII-----EDMNNAI 210			
DB	234	PPGTGLHRSLOAKKMMVQVERIIIEGKIRTKNKREEDVIADVVDVLLKDSSEHLTNL 293			
QY	211	REEDLDSIISNED-----EENAIIRA-KKGDELLNMEDY 245			
DB	294	IANNIDIMMIRGHDSVPVLTILAVKFLSDSPALNLLTEENNKLSKLEELGEPYWNLY 353			
QY	246	QKMETQCVISBALRCGNIIVTKHRAKANDIDKEFELYVPRKGMKVPPIPTAVLADPSLHN 305			
DB	354	LSLPTQKVIETELRMGNVITIGVMKKAKVDVIEKGYVTPKWCFLAYLRSVHLKLYVES 413			
QY	306	PPEFPRMNV-----TKTTAFGGGVRCVPGSELCKLDIAFLIHLNVLSTYR 349			
DB	414	PKFPRMNVQERDMNTSSFSPPGGGQRLCPGLDLARLETSVFLNHLVRRFR 464			

	RESULT	9	
0940V4			
ID	Q940V4	PRELIMINARY:	PRT: 465 AA.
DC	Q940V4;		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DI	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DE	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	ATG3G30180/T20F20_6.		
OS	Aralbidopsids thalianae (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
CC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
CC	eumustoids II; Brassicales; Brassicaceae; Arabidopsids.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Chauk R., Chen H., Kim C.-J., Koesema E., Meyers M.C., Banh J.,		
RA	Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,		
RA	Ishida J., Jiang P.X., Jones T., Kamita A., Kerlin-Neumann G.,		
RA	Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.		
RA	Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.		
RA	Saitou M., Seki M., Southwick A., Tang C.C., Tortumi M., Yamada K.,		
RA	Yanamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,		
RA	Ecker J.R.;		
RT	"Arabidopsis cDNA clones.";		
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBD databases.		
DR	EMBL: AY052655; AAK96559.1; .		
SO	SEQUENCE 465 AA; 53814 MW; 280A2D0172FCall CRC64;		

Query Match	28.2%	Score 572:	DB 10:	Length 465:
Best Local Similarity	30.7%	Pred. No. 1e-37:		
Matches 136:	Conservative 71:	Mismatches 152:	Indels 84:	Gaps 9:
QY	1	MGWPIGTETISFFKPHRSDSITGFLQQRYSRYGKYVKSNICGGKAVVSCDQELNFIION	60	
DB	39	MGWPIGTETTELKQGPDP-----FMKNQRLRGSEFFKSHILCCPTVSDMADELNRILNN	93	
QY	61	EGLKFTSDPKAMHDILGKSYLLATGGEIHRKLKNVIFINLTKSKPDQLCAENLSTS	120	
DB	94	ESKGLVAGI PQGMIDLIGICNIAVHGPRHLMRQSLSLISPTMKDHLLEKIDDFMKN	153	
QY	121	ILKSKNCKEVEFEHKVMMFTLSVYNOLLST-----KPE-DPARLYLQDLFLSTWKGFI	174	
DB	154	YLCGMDDELETYDIOEKTKMP--APLSLQLQIAETLTKKEVEEYR---TEFFKLVVGLT	206	
QY	175	SLPIPLPGGYNALKNVSNRNIIHQNALIEDNNNAIREE-----DELDSIISNED----	224	
DB	207	SVPIIDIPGNIYNSGVQARNRINDRLLELMOE-----RKSEGETFTDMGLYLLKKEDNRLL	261	
QY	225	-----EENHAIRAKGGGELLN	241	
DB	262	LTDKEIRDOVVTILYSGYETVSTSMALMLYLDHPRKALEELREHLAIREKKRPDEPLT	321	
QY	242	WEDYQKMEFTQVISEALRCSNIYVTHRKATADIKFEYVLPKMGKVPPIPTAVHLDS	301	
DB	322	LDDISMKFTRAVITETSRLATIVNGVLKRTTHDELNGYLLPCKGRITVYREIYDTS	381	
QY	302	LHNEPFEENPMKWTKT-----AFGGVYVPCGEGELGKLQIAFFLNLHLVLSRWKIK	353	
DB	382	LYEDMIFENPMMWMEKKSLESKSYFLFEGGVALCPGKELGISEVSSFLHYPTVKRWMEEN	441	
QY	354	SDENPIAHRYVEFKRGMLLEIEP	376	
DB	442	GEDKLMEFPRVYSAAPKHLKSP	464	
RESULT	10			
Q9LN73				
AC	Q9LN73:	PRELIMINARY:	PRT:	478 AA.
DT	01-OCT-2000	(TREMBLrel. 15, Created)		
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)		
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)		

DE CYTOCHROME P450 (BRASSINOSTEROID-6-OXIDASE).
 GN BR60X1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=98290546; PubMed=9628582;
 RA Sato S., Kaneo T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
 RA Tabata S.,
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
 RT Sequence features of the regions of 1,456,315 bp covered by nineteen
 RT physically assigned P1 and TAC clones.";
 RL DNA Res. 5:41-54(1998).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=21295370; PubMed=11402205;
 RA Shimada Y., Fujioke S., Miyauchi N., Kushihiro M., Takatsuto S.,
 RA Nomura T., Yokota T., Kamiya Y., Bishop G., Yoshida S.,
 RT "Brassinosteroid-6-oxidases from Arabidopsis and tomato catalyze
 RT multiple C-6 oxidations in brassinosteroid biosynthesis.";
 RL Plant Physiol. 126:770-779(2001).
 CC 1 - SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AB009048; BAB08653.1; -
 DR EMBL; AB035868; BAB08581.1; -
 DR InterPro; IPR001128; CYP_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
 DR Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 465 AA; 53767 MW; 0C00459C9C866D1F CRC64;

Query Match 26.9%; Score 546; DB 10; Length 465;
 Best local similarity 29.7%; Pred. No. 1.3e-35;
 Matches 130; Conservative 73; Mismatches 161; Indels 74; Gaps 9;

QY 1 MGMPFGETISFEKPHRSDSIGTFLQORVSRGKFKSNICGGKAVVSCDQELMPLFION 60
 DB 39 MGMPFGETFEELKQGRN-----FMNRQRLAYGSEFKSHLLGCPPLISMDESVNRITLKN 93
 QY 61 ESKLETSYPRKAMNDILGKISLLATGEIRKLNVIISFTNLTKSPDFLHCAENLSIS 120
 DB 94 ESKGLVPGYPOSMDILGTCNMAAVHGSSHRLMGSLSLISSTMRRDHLPRKVDHEWRS 153
 QY 121 ILKSNKRCREVEFEHKEVMEFTLSVNVNQLS--IKRDPARLYVQ--DFLSYMKGFIS 175
 DB 154 YLDQWNELEVIDIDQKTHMAFLSSLTQIAGNLKRP-----FVEERKTAFLKLVGSLTS 207
 QY 176 LPIPLGCTGYTNAIKVRSNRNIHONAIIEDMNA-----IREDEF 215
 DB 208 VPIDLPNTNRCGQIARNNIDRLRELMQERRDSGETFDMLGYLMMKEGRIYLTDEBI 267
 QY 216 LDSIIS-----NEDEHNAITAKGDELLMWEYDQK 247
 DB 268 RDOVVTILISYETVSTSMALKYLHDHPKALQELRAEHLAFERRKRODEPLLEDEYKS 327
 QY 248 MEPTQVISEALRCGNIVKTVHRRKATHDIKFEKVIYPRKMKVPFIPFAVHLDPSLHNP 307
 DB 328 MKFRAVIYETSRLATIVNGRLKTRDLEINGLIRKGRIVYTRTEINDANLYEDPL 387
 QY 308 EENMRRTKTT-----AFGGVAVCPGEGELGKQIAFLHLHLVLSYRW-KIKSDMP 358
 DB 388 IFNEMRMKMSLESQNSCFVGGGTGLCPGKEGLIVISSFLHYFYTRYRMEELGDEL- 446
 QY 359 IANPYEFKRGMLLEIP 376
 DB 447 MVPRVFAVAPKGFHLRISP 464

RESULT 13
 ID 0949P1 PRELIMINARY; PRT: 467 AA.
 AC 0949P1;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE PUTATIVE CYTOCHROME P450 PROTEIN.
 GN AT4G19230.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Pham P.K., Banh J., Banno F., Dale J.M.,
 RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Tortum M., Yamamura Y., Yu G., Yu S., Bowser L.,
 RA Carinici P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kanilya A., Karlin-Neumann G., Kawai J., Kim C., Koeseema E., Lam B.,
 RA Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shim P., Southwick A., Tracy S.E.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,
 RT "Full length cDNA of gene At4g19230 (GI:7268718)."
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AT050980; AAR93657.1; -
 SQ SEQUENCE 467 AA; 53037 MW; 2F4230446536D955 CRC64;

Query Match 25.3%; Score 513.5; DB 10; Length 467;
 Best local similarity 29.4%; Pred. No. 5.1e-33;
 Matches 120; Conservative 69; Mismatches 150; Indels 83; Gaps 13;

QY 1 MGMPFGETISFEKPHRSDSIGTFLQORVSRGKFKSNICGGKAVVSCDQELMPLFION 60
 DB 41 MGMPVEGETFQLY-----SDPNVFEQSKKRYGSVFETHVLGCPVWISSPEAKFVLYT 96
 QY 61 ESKLETSYPRKAMNDILGKISLLATGEIRKLNVIISFTNLTKSPDFLHCAENLSIS 117
 DB 97 KSHLEKTFEPASKERMKGQAIFFHOGDYHAKRLKVLRAFMPEESIRNVPDI-----EST 152
 QY 118 SISLSKMKRCREVEFEHKEVMEFTLSVNVNQLSI-----KRPDPARLYVQDFLSY 170
 DB 153 AQSLSRWSESTM-INTQEMKTYTFNV--ALLSTGKDEVLYREDKRCYTLLE----- 203
 QY 171 KGFSLPIPLGCTGYTNAIKVRSN-RNIHONAIIEDMNAIREDDEFDSITSNED----- 224
 DB 204 KGYNSMPVNLPGTLFHKSMKARELSQILARILSERONSSHNDLGSFGMDKEELTDE 263
 QY 225 -----BEHNAITAKGDELLMWEYDQK 245
 DB 264 QINDNITGVFAARDTASVSWILKYLAENPNVLEAVTEEMAKIDKEEGSLTWGDT 323
 QY 246 QKMEFTQCVISEALRCGNIVKTVHRRKATHDIKFEKVIYPRKMKVPFIPFAVHLDPSLHNP 305
 DB 324 KKMPLTSRVYQETLRVASILSFFREAVDEYEGELIRKGRIVYTRTEINDANLYEDPL 383
 QY 306 PEFENMRW-----TKTAFGGVAVCPGEGELGKQIAFLHLHLVLSYRWK-KSDM 357
 DB 384 PGKEDFSREVAKRPMTPEFGNGTHSCPGNELAKLEMSIMIHLLTTKYSNSIVGASDGI 443
 QY 358 ---PIANP 362
 DB 444 QYGFALP 451

RESULT 14
 ID 09FT38 PRELIMINARY; PRT: 518 AA.
 AC 09FT38;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)

30/08:40:07 2002

us-09-995-917a-1.rspt

Page 9

THIS PAGE BLANK (USPTO)

COMMENT

Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bbsrc.ac.uk
Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.

FEATURES

Location/Qualifiers

1..100269

/organism="Arabidopsis thaliana"

/variety="Columbia"

/db_xref="taxon:3702"

/chromosome="5"

2255..6118

/gene="F18022_10"

complement(join(2255..2376,2496..2602,2699..2773,

2896..3004,3094..3317,3419..3493,3582..3743,3871..4005,

4364..4393,4488..4553,4723..4759,4849..4922,5124..5180,

5289..5339,5466..5531,5640..5775,5930..5981,6091..6118))

/gene="F18022_10"

complement(2255..2376)

/gene="F18022_10"

/number=1

complement(join(2255..2376,2496..2602,2699..2773,

2896..3004,3094..3317,3419..3493,3582..3743,3871..4005,

4364..4393,4488..4553,4723..4759,4849..4922,5124..5180,

5289..5339,5466..5531,5640..5775,5930..5981,6091..6118))

/gene="F18022_10"

/note="strong similarity to protoporphyrinogen IX oxidase, glycine max, Ab025102"

/codon_start=1

/product="protoporphyrinogen oxidase-like protein"

/protein_id="CAB87761.1"

/db_xref="GI:7573447"

/translation="MASGAVADHOTEAVSGKRVAVVAGVSGSLAAYLKSRGLNVTY

FEADGRVGGKLRVYQNGILNDEBANTMTBAPEVGSILDDGLREKQFPISQKRY

IVRNGVPVLMPLNPLELTVTSVLTQSFKQILLEPFLMKKSSKVSASAEVSVEFF

QRHQEVVDYLDIDPEVGTSAADPDSLMSKHSFPDLNNSFGSIIVGAIKFKFARGG

KSRDTSKSPGTRKSGSFRGGMQILPDLCKLSHDEINLDSKVLISYNSGRQ

ENMSLSGVSHNETORONPHYDAAPLCNFKEMKVMKGQGPOLNPLPEINMPLSVLIT

TPTKEKVRPLGGFVLIPSKRQKGFETGLPLSSMMFPDRSPEDVHLTYTFGGSR

NOELAKASTDLEKQVYTSDLRLGVEBEPESVNHVYMKRAFPDYDSSYDSGMEYIDK

MENLPGFFVAGNERGSLVSGKSLASGCKADLVISTLSCSNDKKPMDSL"

complement(2377..2495)

/gene="F18022_10"

/number=1

complement(2496..2602)

/gene="F18022_10"

/number=2

complement(2603..2698)

/gene="F18022_10"

/number=2

complement(2699..2773)

/gene="F18022_10"

/number=3

complement(2774..2895)

/gene="F18022_10"

/number=3

complement(2896..3004)

/gene="F18022_10"

/number=4

complement(3005..3093)

/gene="F18022_10"

/number=4

complement(3094..3217)

/gene="F18022_10"

/number=5

complement(3218..3418)

/gene="F18022_10"

/number=5

complement(3419..3493)

/gene="F18022_10"

/number=6

complement(3494..3581)


```
/protein_id="AAL06567.1"
/db_xref="GI:15724348"
/translation="MEEETHHTLLPLLLPSLISLILFLIKRNRKRNLPSPGKS
GMPRLGETIGYIKPYTATTLGDPMOOHVSKYKITYRSNIFGPPTIVSADAGINRITLO
NBRLEPCSTPRISIGILGKMSMLVLDMDHMDMSISLNFISHARLRTILKDVBRH
TLEVLDSWQONSIFSAODEAKKFTFNLMAKHIMSDPGEETEOUKKEVITMKGVVS
APLNPGTAHYKHALOSRATILKIERKMERLIDKKEDEOEVEKTEDEAMSKSDH
VKORITDDLLGWLKHSNLSTEOILDLISLTFAGHETSSVAIALIEPFIQACRAV
BELREHEIARAKKEIGSELMDMDYKMDPTOCVINFTELVKGVVRIHRAIKADY
RYKGDIPSGMKVLPVISAVIDNSRYOQPNFMRMOQOONNGASSSGSFSFTWGN
NMPPGGGPRLCASELKLKEMAVFIHLVLKFNMLADBNPFAPFPDFPNGLPIR
VSRIL"
3'UTR 1571..1691
BASE COUNT 524 a 286 c 363 g 518 t
ORIGIN

Query Match 15.2%; Score 174.6; DB 8; Length 1691;
Best Local Similarity 53.6%; Pred. No. 8.1e-31;
Matches 363; Conservative 0; Mismatches 314; Indels 0; Gaps 0;

QY 4 gtagggccttcattggaaaactatttcttcttcaaacctcatgatcagaactcacc 63
DB 161 GGTGGCCATTCTTGCTGGTAAACCATCGGTATCTTAAACCGTACACGCCACACACATC 220
QY 64 ggtacatcttcgaacaacagtggttcacglatggaagaagtggtcaagtcacaatatatgt 123
DB 221 GGTGATCTTATGCAACACATATGCTCCAGTATGATGATATATATGATGCAACTGTTT 280
QY 124 ggtggaagaagcagtagctcatcgtgacgaagaactcaacatggttcaactcaaaagaa 183
DB 281 GGGAACACACACATCTCATGCTAGCTGATGCTGAATATGATATGATATCAATAACGAA 340
QY 184 gggaaagtgtttacacgattatccaaagaagcagtcgaacttcgtggcaaatcttc 243
DB 341 GGAAGGCTCTTGGAATGTATCTTAAAGTATGAGTGGGATCTTGGGAATGCTG 400
QY 244 ctctctatagccacccggaagaattcacaagaactcaaaaatgatatatagctcatc 303
DB 401 ATGCTTGTCTTGTTGGTACATGATAGATAGATAGATAGATATCTCGCTTAACCTCTTA 460
QY 304 aatctcaacaagtcgaacacttcttcacgtcgcaagaagaacactctctatctcgata 363
DB 461 AGCTACGACGCTTGAACCTATTTACTTAAGATGTTGAGAGACATCTTGTGTTGTT 520
QY 364 ctaaaagtcatggaagaatgctcgagaagtcgaattccatcaagaagttaaagtttact 423
DB 521 CTTGATCTTGGCACAAACTATATTTCTCTGCTCAAGACGAGGCCAANAAGTTTACG 580
QY 424 ctcaagtgtatgtaaaccaactcttgagcatcaagccagaagaacccagaagacttata 483
DB 581 TTTAATCTAATGGCGAAGCATATATAGATAGATAGATCTGAGAAAGAACAGAGCA 640
QY 484 gtaatgcaagaatcttctatctatagaaggttatctccttaccataccgctcca 543
DB 641 TTAAGAAAGAGTATGTATCTTCAATGAAGAGTGTCTGCTCTCTTAATCTACCA 700
QY 544 ggaacgggtgtatacaaacgcaatlaagtgtagatccaatccgtaataatatacaaacgca 603
DB 701 GGAACGCTTATCATTAAGCTCTTCAAGTACAGACACATATTTGAATTCATTGAGAG 760
QY 604 attatagaagacatlaataatgaataagaagaagatttctgattcgataatcgc 663
DB 761 AAAATGGAAGAGAGAAATTTGATATCAAGAGAAAGATCAAGAGCAAGAGAGAGTGA 820
QY 664 aatgaagaatgaagaaca 680
DB 821 ACAGAGAGATGAAGACAGA 837

LOCUS 3
AV063728 AY063728 1398 bp mRNA linear PLN 05-DEC-2001
```

```
DEFINITION
Arabidopsis thaliana AT3g30180/T20F20_6 mRNA, complete cds.
ACCESSION
AY063728
VERSION
AY063728.1 GI:17380629
KEYWORDS
FLI_CDNA.
SOURCE
thale cress.
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1398)
REFERENCE
Kim,C.J., Chen,H., Cheuk,R., Meyers,M.C., Banh,J., Bowser,L.,
Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
Ondera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
and Ecker,J.R.
Arabidopsis ORF clones
Unpublished
2 (bases 1 to 1398)
TITLE
Direct Submission
JOURNAL
Submitted (15-NOV-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
COMMENT
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : "RIKEN
Arabidopsis Full-length cDNA") : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Kim,C.J., Chen,H.,
Cheuk,R., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Chang,E.,
Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B.,
Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Ondera,C.S., Palm,C.J.,
Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A.,
and Ecker,J.R.
Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.
FEATURES
source
location/Qualifiers
1..1398
/oranism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="3"
/clone="U13734"
/note="this clone is in pUNT 51
ecotype: Columbia"
1..1398
/note="putative cytochrome P450 homolog"
/codon_start=1
/product="AT3g30180/T20F20_6"
/protein_id="AAL36078.1"
/db_xref="GI:17380630"
/translation="MGIMAMILGLIVICLTALLRNQMRYSKSGLPPTMGPIF
GETTEFLKOGDEPMKORLRYGSEPKSHILCPTIVSDAENRITILNNEKSGVAGY
POSMIDITGTCNIAAYVGHGSPHRIMGSLISLIPMTMCDHLPLKIDPMRMYIGWMD
LETVDIOKTRHMAFLSSLDIOAETLKPEVEYERTEFKLIVGLTSPIDIPGNIR
SGVOARNIDRLTELEWDERKESGETFDMDGLMKEDKNRYLLTDKEIRQDVITILY
SGEIVSTSMALKYLDHDPKALEELREHAIERERKRPDEPLTDIDIKMKETTRAV
IETSRLATIVGVLRKTTHDLELNGYLIPIKGMRIYVYTRFINDTSLYEDPMIFNPV
```

```

RMKESLSKSYFLRGGGVRCLCPGKELGISEVSSFLHYFTVTKRMENGDKLWFP
RVSAPEGYHLKCSpy"
BASE COUNT      448 a      258 c      304 g      388 t
ORIGIN

Query Match      9.4%; Score 107.6; DB 8; Length 1398;
Best Local Similarity 56.1%; Pred. No. 4.8e-15;
Matches 240; Conservative 0; Mismatches 164; Indels 24; Gaps 1;

Oy 675 agaacatgcagccattagagccaagaaggagggaggaacttttgaaattggaagatta 734
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 915 AGAACATTTGGCTTAAAGGGAGAGAAACGACCTGACACCCCTCCTCGACGATAT 974

Oy 735 tcagaagatggaattcactcaatgtgtgatttcgaaggcactacgaatggttaataatcgt 794
    || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 975 TAAATCGATGAATTTCACTCGACGTGATCTTTGAGACATCAAGATTGGACAGATTGT 1034

Oy 795 caagaactgatacagaagaagcactacatgataataaattcaagaagaatgtaaccaa 854
    || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1035 TAAAGTGCTCTTACGAAACACTACTACGACTTAGAACCTCAACGGTTATTAAATCCCAA 1094

Oy 855 ggggtggaagggtgttcccaactcacaagcagatacatttgatccctctctcatgaaaa 914
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1095 AGGTTGGAAATTCATGATACACAGAGATTAACTATGATATCTTTATGAAGA 1154

Oy 915 tcccttgaattaatcccaagagatgagcaaaaaga----- 952
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1155 TCCAATGATCTTTAAACCATGAGATGATGAAAAGACTTGAATCAAGACTATT 1214

Oy 953 --cggcgttggaggagaglaaggatgtctcctgtgtggaacttggcaagctccaaat 1010
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1215 CTTACTCTTTGGAGTGAGTGAAGCTTTGCCCTGGAAGAGTAAGGAATCTCGAAGT 1274

Oy 1011 tgccttctctctcatcactctgtccctcctcatatggtggaataaagaatcagaatgaat 1070
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1275 CTCAGACTCTCTTACCTCTTTTACAAAAATATGATGAGAGAGATGGAAGACAA 1334

Oy 1071 gccaatgc 1078
    ||| |
Db 1335 ATTAAATGC 1342

RESULT 4
LOCUS      AY052655      1722 bp      mRNA      linear      PLN 05-SEP-2001
DEFINITION Arabidopsis thaliana AT3g30180/720P20_6 mRNA, complete cds.
ACCESSION  AY052655
VERSION     AY052655.1 GI:15450574
KEYWORDS   F11-CDNA.
SOURCE     thale cress.
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE  1 (bases 1 to 1722)
            Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
            Bower,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawal,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
            Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
            Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
            Taniuchi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
            Davis,R.W., Theologis,A. and Ecker,J.R.
            Arabidopsis cDNA clones
TITLE      Unpublished
JOURNAL    2 (bases 1 to 1722)
AUTHORS    Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
            Bower,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawal,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
            Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
            Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
            Taniuchi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
            Davis,R.W., Theologis,A. and Ecker,J.R.

```

```

TITLE      Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
            Davis,R.W., Theologis,A. and Ecker,J.R.
JOURNAL    Direct Submission
            Submitted (17-AUG-2001) Salk Institute Genomic Analysis Laboratory
            (SIGNAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT    RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawal,J.,
            Hayashizaki,Y. and Shinozaki,K.

FEATURES
    source
        location/Qualifiers
            1..1722
            /organism="Arabidopsis thaliana"
            /db_xref="taxon:3702"
            /chromosome="3"
            /clone="RAF106-74-K19(R1374)"
            /note="ecotype: Columbia"
            1..91
            92..1489
            /note="putative cytochrome P450 homolog"
            /codon_start=1
            /product="AT3g30180/720P20_6"
            /protein_id="AAK96559.1"
            /db_xref="GI:15450575"
            /translation="MGIMMMITGLVITVICTALLRMNOMRYSKKGPPCTMGPIF
            GEMPEFLKQGPDMKRNRLRYGSFPRKHIIGCPITVIMDAELNRYLIMNESKGLVAGY
            PQSMIDLTCNTIAAVRGPRIKLRGSLISISTPMKDHLLPRIDPFMRNYLCQMD
            LETVIDEKTNNHAFVSSLDIAETLKPVEEYRTFEFKLVGTLSVIDIPGTRR
            SGVARNRIDRLLELMQERKESGEFTDMGVLMKREDNRVLTDEIIRVVYTLIY
            SGYEIVSTTSMALKYLRDHPKALEELREHLAIREKRKRPDEPLTDDIKMKFTFV
            IFEFSRLATYNGVLRKTHDLELNGVILYIGMRIVYTRERINDTSLYEDPMLENFV
            RMKESLSKSYFLRGGGVRCLCPGKELGISEVSSFLHYFTVTKRMENGDKLWFP
            RVSAPEGYHLKCSpy"
            1490..1722
BASE COUNT      585 a      301 c      351 g      485 t
ORIGIN

Query Match      9.4%; Score 107.6; DB 8; Length 1722;
Best Local Similarity 56.1%; Pred. No. 4.7e-15;
Matches 240; Conservative 0; Mismatches 164; Indels 24; Gaps 1;

Oy 675 agaacatgcagccattagagccaagaaggagggaggaacttttgaaattggaagatta 734
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1006 AGAACATTTGGCTTAAAGGGAGAGAAACGACCTGACACCCCTCCTCGACGATAT 1065

Oy 735 tcagaagatggaattcactcaatgtgtgatttcgaaggcactacgaatggttaataatcgt 794
    || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1066 TAAATCGATGAATTTCACTCGACGTGATCTTTGAGACATCAAGATTGGACAGATTGT 1125

Oy 795 caagaactgatacagaagaagcactacatgataataaattcaagaagaatgtaaccaa 854
    || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1126 TAAATCGATTTCTTACGAAACACTACTACGACTTAGAACCTCAACGGTTATTAAATCCCAA 1185

Oy 855 ggggtggaagggtgttcccaactcacaagcagatacatttgatccctctctcatgaaaa 914
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1186 AGGTTGGAAATTCATGATACACAGAGATTAACTATGATATCTTTATGAAGA 1245

```

Query Match	8.6%	Score 98.8	DB 8	Length 1741
Best Local Similarity	52.7%	Pred. No. 5.5e-13		
Matches 257	Conservative	0	Mismatches 207	Indels 24
				Gaps 1

OY	676	gaacatcgaccatttagagccaagaaggagtgtgggaactttigaattggaagttat	735
Db	1027	GAGCATTTCGCAATTCACGGAAAAAACAACAGACGAACCACTCGGTCTTTAGGCGCTG	1086
OY	736	cagaagatgaaatcaactcaaatgttgatcttcaggacctacgaatgtgatataatcgtc	795
Db	1087	AAGTCAATGAAGTTCACTGCAGCTGTGATTTATGAGACATCAAGATTGGCAACGATCTT	1146
OY	796	aagactgtcatagaaaagctcactatgatatattaattcaagaataatgtatccaaag	855
Db	1147	AATGGGCTCTTAAGGAAAACACTCGTGACTTTGAAATCAACGGTATTTAATCCAAAA	1206
OY	856	gggttgaagtggttccaaacttcacaagaagfacatcttgaccccccttcatagaat	915
Db	1207	GGATGAGCAATTTATGCTATACACGAGGGAATAATTAATTACGATGCCAAATCTTATGAAAC	1266
OY	916	ccctttaattaatccatccatgagatgagaccaaacga-----	952
Db	1267	CCATTGATCTTTAATTCACATGAGATGAGATGAGAAGAGAGCTTGAGATCACAAATCATCG	1326
OY	953	-cggcgtttggaggaggagtaaaggatgtcctgtgtgtgtaacttggcaagctccaatt	1011
Db	1327	TTTTGTTTGGAGGGGCGACCAAGGCTTTGCTCGTGTAAGCAACTAGCGATTTGTGCGATC	1386
OY	1012	gctcttcctccatcatctgtctcctcctctaagtgtgaaaaataaagcacatgaatg	1071
Db	1387	TGCGACTTCTTCATTAATCTTCTTTACGAGATACAGATGGGAGAGAAATGAGAGGGATGAA	1446
OY	1072	ccaatcgcgacccttaactgcgttgagattaagaagaagcatgctttggagataagccaaca	1131
Db	1447	TTAAATGCTGTTTCCAGAGATTTTTCACCACMAAGCTTCATCTTAGCATTCACCCCTAC	1506
OY	1132	aaattcct 1139	
Db	1507	TTTTTCT 1514	
RESULT	6		
LOCUS	AY063722	1419 bp mRNA linear PLN 05-DEC-2001	
DEFINITION	Arabidopsis thaliana AF590569/MJ03_9 mRNA, complete cds.		
ACCESSION	AY063722		
VERSION	AY063722.1 GI:17380617		
KEYWORDS	FLI.CDNA.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi		
AUTHORS	1 (bases 1 to 1419) Kim,C.-J., Chen,H., Cheuk,R., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Oondera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.		
TITLE	Arabidopsis ORF clones		
JOURNAL	unpublished		
REFERENCE	2 (bases 1 to 1419)		
AUTHORS	Kim,C.-J., Chen,H., Cheuk,R., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Oondera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-NOV-2001) Salk Institute Genomic Analysis Laboratory		
Journal	(Sigmbl), Plant Biology Laboratory, The Salk Institute for		

Db 933 AGGAGAGCATGAAAGATTAGGCAATGAGAGTATTCGTATAGTCTTGAATGAGTG 992
Qy 731 attatcagaagaatggaatcactcaatgctgattctcgaagcactacgactgtaata 790
Db 993 ATTACAAATGATCATGCCATTGCACACATGTGTATATGAGACGCTACGAGTGGCTTACA 1052
Qy 791 tgcgaagcctgacatagaaaagctaccatcatatataatcaagaatctgattc 850
Db 1053 TCATCGCGCGGTGTTTTCACACAGCGCAATGACGATTTAGATCAAAAGCTTATAAATTG 1112
Qy 851 caaagggtggaaggtgcttcacatctcagcagcagatcattgactccctctcatg 910
Db 1113 CAAAGGCGGGAAGATATCTCATCTGTAGAGCGGTGCTTATAGACCAACCAACTTCA 1172
Qy 911 aaaaatccttgaaattcaatcccatgagatgagccaaacagcg 955
Db 1173 AAGATGCTCGCACTTCAACCCCTTGAGATGGCAGCAACTCGG 1217

RESULT 10
AY042837 1641 bp mRNA linear PLN 04-JUL-2001
LOCUS Arabidopsis thaliana cytochrome P450 90A1 (MJJ3.9) mRNA, complete
DEFINITION cds.
ACCESSION AY042837
VERSION AY042837.1 GI:14596098
KEYWORDS FLI.CDNA.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1641)
Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,
Palm,C.J., Bower,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Chen,K., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinozaki,Y., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submision
Submitted (25-JUN-2001) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
COMMENT e-mail for correspondence: arabsequence.stanford.edu

TITLE
JOURNAL
Riken Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : RIKEN
Arabidopsis Full-Length cDNA); Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

FEATURES
source
1.1641
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="V"
/clone="RAF104-19-J21"
/note="This clone is in pluscript
ecotype: Columbia"
1.1641
/gene="MJJ3.9"
20.1438
/gene="MJJ3.9"

/codon_start=1
/product="cytochrome P450 90A1"
/protein_id="AAK68777.1"
/db_xref="GI:14596098"
/translation="MAATFAFLLLSLSTAGFLLLRLRRYRMRKLPKSGSLPLIGET
FOLIGAVKTEPPEPIDEVARVGSVETTHLEGETTISADPETNRVYLOWEGLFEC
SYPSICNLCKSHLLMKSGILHRMHSLSFNSSILKHLMLDIDRLVRLMDM
SSRYLMEAKKITFEELVKOLMSDEGEMESLRKEVLVIEGFSLPLFSTYR
KAIGARRVAVALVVMVKRRREEEGEGRKMLAALLAADGSDDEEIVDFVAL
VAGETSTMTLAVKLTETPLALAEKHEKIRAKMSDSYSLEMSDYSMEFTOC
VNETLRANITIGVPRAMTDVIEKIKTKGKVSSEFAVHLDPHDFDARTFN
WRWQSNSTVTPSPVETFPFGGPRICPEYELARVALSVFLRLVTFGSWVPAEDKLV
FFPTTRQKRPVIEVKRDFAT"

BASE COUNT 402 a 371 c 380 g 488 t
ORIGIN

Query Match 8.5%; Score 97.8; DB 8; Length 1641;
Best Local Similarity 58.9%; Pred. No. 9,5e-13;
Matches 168; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

Qy 671 atgaagaacatgcaagccatagagcaagaaggagatggaacttgattggaag 730
Db 933 AGGAGAGCATGAAAGATTAGGCAATGAGAGTATTCGTATAGTCTTGAATGAGTG 992
Qy 731 attatcagaagaatggaatcactcaatgctgattctcgaagcactacgactgtaata 790
Db 993 ATTACAAATGATCATGCCATTGCACACATGTGTATATGAGACGCTACGAGTGGCTTACA 1052
Qy 791 tgcgaagcctgacatagaaaagctaccatcatatataatcaagaatctgattc 850
Db 1053 TCATCGCGCGGTGTTTTCACACAGCGCAATGACGATTTAGATCAAAAGCTTATAAATTG 1112
Qy 851 caaagggtggaaggtgcttcacatctcagcagcagatcattgactccctctcatg 910
Db 1113 CAAAGGCGGGAAGATATCTCATCTGTAGAGCGGTGCTTATAGACCAACCAACTTCA 1172
Qy 911 aaaaatccttgaaattcaatcccatgagatgagccaaacagcg 955
Db 1173 AAGATGCTCGCACTTCAACCCCTTGAGATGGCAGCAACTCGG 1217

RESULT 11
AB066286 1476 bp mRNA linear PLN 20-JUL-2001
LOCUS Arabidopsis thaliana mRNA for CYP90D, complete cds.
DEFINITION Arabidopsis thaliana mRNA for CYP90D, complete cds.
ACCESSION AB066286
VERSION AB066286.1 GI:14971016
KEYWORDS
SOURCE
Arabidopsis thaliana (cultivar:Columbia) cDNA to mRNA,
clone:PGMD26.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1476)
Shimada,Y.
P450 gene repressed by brassinosteroid
Unpublished
2 (bases 1 to 1476)
Shimada,Y.
Direct Submision
Submitted (18-JUL-2001) Yukihisa Shimada, RIKEN, Plant Science
Center, 2-1, Hirosewa, Wako, Saitama 351-0198, Japan
(E-mail:shimada@postman.riken.go.jp, Tel:81-48-467-9529,
Fax:81-48-462-4959)

FEATURES
source
1.1476
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/clone="PGMD26"
/note="P450"

		/note="compared to genomic sequence"			
		/replace="ttg"			
BASE COUNT	627 a	398 c	387 g	596 t	
ORIGIN					
Query Match	7.9%; Score 90.4; DB 8; Length 2008;				
Best Local Similarity	54.2%; Pred. No. 5.1e-11;				
Matches 218; Conservative	0; Mismatches 166; Indels 18; Gaps 1;				
QY	672	tgaagaacatcagccattagagccaagaaggagatgggaactttggaattgggaaga	731		
DB	1051	TGAAGAACAAATGGCAATTAAGCAAGACAAAGAGAGTCTTCACTTGGGGAGA	1110		
QY	732	ttatcagaagatggaattcactcaatgctgctgattcttcagagcaactggtgtaaat	791		
DB	1111	TACAAAGAAAGATGCCATTAACTTCAAGAGTTTATCAAGAACATTTAAGAGTCGCTTCAAT	1170		
QY	792	cgtcaagactgacatagaaaagctactcatalataataatcaagaatatgattcc	851		
DB	1171	CTTATCTTTACATTCAGAGAAAGCTGCGAAGATGTGCAATACGAAGATATTGTATAC	1230		
QY	852	aaaggggaggaagggtgttccaaattccacagcagatattctgtccctctcaca	911		
DB	1231	TAAAGGATGGAAAGTGTACCCCTTATCGAAMAACATTCATGCTGATATTTTTTC	1290		
QY	912	aaatccttgaatttaattccatgagat-----gagccaaaagcag	953		
DB	1291	TAAATCGGGGAAATTTGATTCATCAAGATTCAGAGTGGCTCCAAANCCCAATAGCTTCAAT	1350		
QY	954	ggcgttggaggaaggagtaagggtatgttcctggtggtgaacttggcaagctccaaatgcg	1013		
DB	1351	GCCATTTGGCAATGGAGACCCACTCGTGTCTGGAATGAGCAAGCTTGAGATGTCT	1410		
QY	1014	tttctctctatcatctctgtctctctctctatagatggtggaatat	1055		
DB	1411	TATTATGATTCTATCTACCTACCAAGTACAGTTGTTCAT	1452		
RESULT 13					
AY065065	1681 bp mRNA linear PLN 07-JAN-2002				
DEFINITION	Arabidopsis thaliana AT5g45340/K9E15_12 mRNA, complete cds.				
ACCESSION	AY065065				
VERSION	AY065065.1 GI:18086489				
KEYWORDS	FLI CDNA.				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
	Rosidae; eurosida II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1 (bases 1 to 1681)				
AUTHORS	Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J., Bower, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamliya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Saitou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.				
TITLE	Arabidopsis cDNA clones				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1681)				
AUTHORS	Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J., Bower, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamliya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Saitou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.				
TITLE	Direct Submission				
JOURNAL	Submitted (30-NOV-2001) Salk Institute Genomic Analysis Laboratory				

		(SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA			
		RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFLP CDNA (RFLP CDNA : "RIKEN Arabidopsis Full-Length cDNA") : Seki, M., Narusaka, M., Ishida, J., Saitou, M., Kamliya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.			
		The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RFLP CDNA: Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shin, P., Banh, J., Bower, L., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.			
		Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.			
FEATURES		Location/Qualifiers			
source		1..1681			
		/organism="Arabidopsis thaliana"			
		/db_xref="taxon:3702"			
		/chromosome="5"			
		/clone="RAFL09-13-J06(R21133)"			
		/note="ecotype: Columbia"			
		1..82			
5'UTR		83..1474			
CDS		/note="cytochrome P450"			
		/codon_start=1			
		/product="AT5g45340/K9E15_12"			
		/protein_id="AAL57698.1"			
		/db_xref="GI:18086490"			
		/translation="WDPSGLPLTISAALFLCLRLFTNGVRSSSTKLPPLPGMGYP			
		IVGETFQITSDPWFVFAKORRTGSVFKTHVLGQPCVWISSPEAKFLVTKSHLF			
		PTFPASRKMLGKQAIFFHQGDYSKRLKLVLAEMPDAIRNWPPIESIAQESLSW			
		DGTOLNTEQEMKTYTFYFNVALISLKGDEYVREDLRQCYLLEKYNSEMPINLPGLT			
		HKAMKARKELAOILANILSKRQNPSSHTDLGSPMEDACLTDOIADNITGVFAA			
		RDTTASVLTWLTKLADNPVLVAETEOAMIRDKKESGLTMDTKMPLTYRVIO			
		ETLRATILSTFPREAVDEYVEGILLIPKGVLPPLRNHNNADIFEDPKFPDSRF			
		EVARPNTPFMPGSGIHSCPNELAKETISVLHLLTKYRMSTVPSDGIQYGFAL			
		PQNDLPALERRP			
		1475..1681			
3'UTR		1475..1681			
BASE COUNT		556 a	353 c	323 g	449 t
ORIGIN					
Query Match	7.8%; Score 89.8; DB 8; Length 1681;				
Best Local Similarity	58.4%; Pred. No. 7.2e-11;				
Matches 157; Conservative	0; Mismatches 112; Indels 0; Gaps 0;				
QY	672	tgaagaacatcagccattagagccaagaaggagatgggaactttggaattgggaaga	731		
DB	988	TGAAGAACAAATGGCAATTAAGCAAGACAAAGAGTCTTCACTTGGGGAGA	1047		
QY	732	ttatcagaagatggaattcactcaatgctgctgattcttcagagcaactggtgtaaat	791		
DB	1048	TACAAAGAAAGATGCCATTAACTTCAAGAGTTTATCAAGAACATTTAAGAGTCGCTTCAAT	1107		
QY	792	cgtcaagactgacatagaaaagctactcatalataataatcaagaatatgattcc	851		
DB	1108	CTTATCTTTACATTCAGAGAAAGCTGCGAAGATGTGCAATACGAAGATATTGTATAC	1167		
QY	852	aaaggggaggaagggtgttccaaattccacagcagatattctgtccctctcaca	911		
DB	1168	AAAGGATGGAAAGTGTACCCCTTATCGAAMAACATTCATGCTGATATTTTTC	1227		
QY	912	aaatccttgaatttaattccatgagat-----gagccaaaagcag	940		
DB	1228	GGATCGGGGAAATTCGATCGCTCGAGAT	1256		

RESULT 14
OSJN00012/c
LOCUS OSJN00012 139969 bp DNA linear HTG 08-SEP-2001
DEFINITION Oryza sativa chromosome 4 clone OSJNBa0089K21, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
ACCESSION AL606441
VERSION AL606441.1 GI:15552679
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
AUTHORS Fu, G., Wang, S.Y., Ren, S.X., Lv, G., Lin, W., Gu, W.Q., Zhu, G.F.,
Tu, Y.F., Jia, J., Yin, H.F., Zhang, Y., Cai, Z., Chen, J., Kang, H.,
Chen, X.Y., Shao, C.Y., Sun, Y., Hu, Q.P., Zhang, X.L., Zhang, W.,
Wang, L.J., Ding, C.W., Sheng, H.H., Gu, J.L., Chen, S.T., Ni, L.,
Zhu, F.H., Han, B., Peng, Q., Huang, Y.C., Li, Y., Zhu, J.J., Zhao, Q.,
Hu, X., Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L., Weng, Q.J.,
Zhang, L., Lu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.T., Zhang, Y.J., Lu, Y.,
Li, C., Li, T., Zhang, Y., Hu, H., Jia, P.X., Qian, Y.M., Ying, K.,
Zhou, B., Chen, Z.H., Hao, P., Zhang, L., Wu, M., Zhang, R.Q., Guan, J.P.
and Hong, G.F.
Direct Submission
Submitted (28-JUL-2000) Han Bin, National Center for Gene Research,
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
clone: OSJNBa0089K21.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1. 139969
/organism="Oryza sativa"
/variety="Nipponbare"
/sub-species="japonica"
/db_xref="taxon:4530"
/chromosome="4"
/clone="OSJNBa0089K21"
/clone_11b="CUGI-OSJNBa"
BASE COUNT 39133 a 31444 c 30948 g 38417 t 27 others
ORIGIN
Query Match 7.8%; Score 89.2; DB 2; Length 139969;
Best Local Similarity 57.0%; Pred. No. 6.9e-11;
Matches 163; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
OY 92 ggtatggaagaagtgttcacgaatataatgtggtggaagaacagtagctcatgtgacc 151
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2033 ggtatggagggtgtttaagctccatctgttctgcaccccccacacatagctcctgtgacc 1974
OY 152 aagaactcaacatgttcatctacatcaaaaacgaaggaagtggttcatcgcgattatccaa 211
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1973 aggacctgaacacattcattccttcagaaatgagagagcgtgtttcagctcactacccca 1914
OY 212 aagcgatgcatacattctcggcaaatatccctctatattagccacggaagaattcaca 271
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1913 gggccattcattgcattctggcgaagtcctcatgttagtgctcctcagggagacaccaca 1854
OY 272 ggaactaaaaatgttatattatagcttcaatcattccacaagtcgaactgacttc 331
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1853 agagcctcaggaaccttgctttagcactgctcactccacaaagtcgaagccacactacc 1794

OY 332 ttcactgcgcagagaacctctctatctcgataactaaatgaatgaa 377
||| | ||||| | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1793 ttggccacacattgagaaattggcacctgcatttacttgggtcattgcca 1748
RESULT 15
OSJN00023/c
LOCUS OSJN00023 169845 bp DNA linear HTG 11-SEP-2001
DEFINITION Oryza sativa chromosome 4 clone OSJNBa0016002, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
ACCESSION AL606388
VERSION AL606388.1 GI:15594047
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
AUTHORS Fu, G., Wang, S.Y., Ren, S.X., Lv, G., Lin, W., Gu, W.Q., Zhu, G.F.,
Tu, Y.F., Jia, J., Yin, H.F., Zhang, Y., Cai, Z., Chen, J., Kang, H.,
Chen, X.Y., Shao, C.Y., Sun, Y., Hu, Q.P., Zhang, X.L., Zhang, W.,
Wang, L.J., Ding, C.W., Sheng, H.H., Gu, J.L., Chen, S.T., Ni, L.,
Zhu, F.H., Han, B., Peng, Q., Huang, Y.C., Li, Y., Zhu, J.J., Zhao, Q.,
Hu, X., Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L., Weng, Q.J.,
Zhang, L., Lu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.T., Zhang, Y.J., Lu, Y.,
Li, C., Li, T., Zhang, Y., Hu, H., Jia, P.X., Qian, Y.M., Ying, K.,
Zhou, B., Chen, Z.H., Hao, P., Zhang, L., Wu, M., Zhang, R.Q., Guan, J.P.
and Hong, G.F.
Direct Submission
Submitted (28-JUL-2000) Han Bin, National Center for Gene Research,
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
clone: OSJNBa0016002.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1. 169845
/organism="Oryza sativa"
/variety="Nipponbare"
/sub-species="japonica"
/db_xref="taxon:4530"
/chromosome="4"
/clone="OSJNBa0016002"
/clone_11b="CUGI-OSJNBa"
BASE COUNT 47749 a 38566 c 37139 g 46370 t 21 others
ORIGIN
Query Match 7.8%; Score 89.2; DB 2; Length 169845;
Best Local Similarity 57.0%; Pred. No. 6.8e-11;
Matches 163; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
OY 92 ggtatggaagaagtgttcacgaatataatgtggtggaagaacagtagctcatgtgacc 151
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 165315 ggtatggagggtgtttaagctccatctgttctgcaccccccacacatagctcctgtgacc 165256
OY 152 aagaactcaacatgttcatctacatcaaaaacgaaggaagtggttcatcgcgattatccaa 211
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 165255 aggacctgaacacattcattccttcagaaatgagagagcgtgtttcagctcactacccca 165196
OY 212 aagcgatgcatacattctcggcaaatatccctctatattagccacggaagaattcaca 271
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 165195 gggccattcattgcattctggcgaagtcctcatgttagtgctcctcagggagacaccaca 165136

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 13:57:33 ; Search time 13.47 seconds
(without alignments)
1098.060 Million cell updates/sec

Title: US-09-995-917A-1
Perfect score: 2027

Sequence: 1 MGWPFIGFTISFFKPHRSDS.....YVEFKRGMLIEPTKFLD 382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SWISSPROT_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query length	DB ID	Description
1	655	32.3	472 1	C901_ARATH 042569 arabidopsis
2	638	31.5	524 1	C90C_ARATH 09m066 arabidopsis
3	568.5	28.0	464 1	CP85_LYCES 043147 lycopersico
4	432.5	21.3	490 1	C883_ARATH 023051 arabidopsis
5	390.5	19.3	519 1	C881_MAIZE 043246 zea mays (m
6	347.5	17.1	444 1	C120_SYNY3 059990 synchocyst
7	294.5	14.5	492 1	CP26_BRAE F79739 brachydanio
8	292.5	14.4	497 1	CP26_MOUSE 055127 mus musculu
9	280	13.8	492 1	CP26_XENLA 093323 xenopus lae
10	278.5	13.7	512 1	C26A_HUMAN 09m63 homo sapien
11	275	13.6	492 1	CP26_CHICK 09m04 gallus gall
12	267.5	13.2	497 1	CP26_HUMAN 043174 homo sapien
13	230.5	11.4	494 1	CPA5_RAT P20812 rattus norv
14	230.5	11.4	494 1	CPA7_HUMAN P20853 homo sapien
15	223.5	11.0	494 1	CPA4_MOUSE P15392 mus musculu
16	223	11.0	500 1	C912_ARATH 065790 arabidopsis
17	219.5	10.8	494 1	CPA5_MOUSE P20852 mus musculu
18	215.5	10.6	494 1	CPA6_HUMAN P11509 homo sapien
19	215.5	10.6	494 1	CPA8_RABIT 005356 oryctolagus
20	206.5	10.2	494 1	CPAA_RABIT 005555 oryctolagus
21	205.5	10.1	494 1	CPA8_MESAU P24454 mesocricetu
22	204.5	10.1	538 1	CP18_DROME 095078 drosophila
23	199	9.8	492 1	CPAC_MOUSE P56593 mus musculu
24	197	9.7	492 1	CPA1_RAT P11711 rattus norv
25	193	9.5	511 1	C7C4_ARATH 064655 arabidopsis
26	190.5	9.4	492 1	CPBC_RAT P33272 rattus norv
27	190	9.4	528 1	CP51_CANTR P14263 candida tito
28	188.5	9.3	493 1	CPA9_MESAU P24455 mesocricetu
29	188.5	9.3	494 1	CPAD_HUMAN 016596 homo sapien
30	188	9.3	510 1	C755_EUSGR 096418 eustoma gita
31	187.5	9.3	494 1	CPBB_CANFA P24460 canis famli
32	186	9.2	528 1	CP51_CANAL P10613 candida alb
33	185.5	9.2	491 1	CPB4_RABIT P00178 oryctolagus

34	185.5	9.2	500 1	CPJ1_RABIT P52786 oryctolagus
35	185	9.1	492 1	CPA2_RAT P15149 rattus norv
36	185	9.1	504 1	CPB2_HEIAM 027664 helioverpa
37	184.5	9.1	522 1	C822_SOYBN 081972 glycine max
38	183.5	9.1	501 1	CP06_MOUSE 054750 mus musculu
39	182.5	9.0	490 1	CP07_RAT P05179 rattus norv
40	182.5	9.0	518 1	CP3R_ONCMW 042563 oncorhynch
41	182	9.0	508 1	CP48_RAT P24464 rattus norv
42	181.5	9.0	495 1	C4S3_DROME 09xyy0 drosophila
43	181.5	9.0	514 1	CP07_ICYPU 073853 icetulus p
44	181	8.9	500 1	C72V_ARATH 091196 arabidopsis
45	180.5	8.9	491 1	CPB5_RABIT P12789 oryctolagus

ALIGNMENTS

```

RESULT 1
ID C901_ARATH STANDARD; PRT; 472 AA.
AC 042569;
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Cytochrome P450 90A1 (EC 1.14.-.-).
GN CYP90A1 OR CYP90 OR CPD OR AT5G05690 OR MJJ3.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Euphorbia II; Brassicales; Brassicaceae; Arabidopsis.
ON NCBI_TaxID=3702;
RX MEDLINE=96200769; PubMed=8612270;
RC STRAIN=CV. COLUMBIA;
RA Szekeres M., Nemeth K., Koncz-Kalman Z., Mether J., Kauschmann A.,
RA Altman T., Redei G.P., Nagy F., Scheil J., Koncz G.;
RT "Brassinosteroids rescue the deficiency of CYP90, a cytochrome P450,
RT controlling cell elongation and de-etiolation in Arabidopsis.",
RL Cell 85:171-182(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=97471969; PubMed=9330910;
RA Sato S., Kocani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 Mb regions covered by twenty physically assigned
RT clones."
RL DNA Res. 4:215-230(1997).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X87367; CAA60793.1; -
CC EMBL: X87368; CAA60794.1; -
CC EMBL: AB005237; BAB0663.1; -
CC InterPro: IPR001128; Cyt_P450.
CC Pfam: PF00067; P450; 2.
CC PROSITE: PS00086; CYTOCHROME_P450; 1.
CC TRANSFERASE: Monooxygenase; Transmembrane; Heme; Multigene family.
CC TRANSFERASE: 1 21
CC BINDING 418 418 HEME (BY SIMILARITY).
CC SEQUENCE 472 AA; 53785 MW; 41A73F46D64E343F CRC64;

```

Query Match 32.3%; Score 655; DB 1; Length 472;

Best Local Similarity 34.5%; Pred. No. 1.1e-37;
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

QY	1	MGNPFIIGETIIEFNRPHRSDSIIGTFLQARVSNXYGVFNISNIGCGAAVYSCDQELMNFLLQ	60
Dp	36	LGPLPIGTFEOLICAYKATENEPEPIDEWAVRXYGVFTMHLEGPTISADPETRFFVLQ	95
QY	61	EGKLETSDDYPRAMHDILGKYSILLATLGEIRKLNVLNIFINTLKSQPDLFHCANSTIS	120
Dp	96	EGKLECECYPASIONLGLKSHLLMKMSLKHRRHMSLIMSFRANSSITIDHMLDIDRLVRF	155
QY	121	ILKSKNCREVEFHEKVMKFTLSVMYNQLLSIKEDPARLYLODFLSYMGFISLPIPL	180
Dp	156	NLDMSMS--RYLLMEBAKKITFELTVKQMSFDECGEMSE--SLKREYLLVTEGFSPSLPL	212
QY	181	PGGTGTATAIKVRSMRNHQNLI-----EDMNNALIREDFDLSITSNED-----	224
Dp	213	FSTTYRKAIOAR--RKVAEALTVVVMKRREEBEGGARKKQMLAALAAADGSEDEIWD	270
QY	225	-----EEHAALRAKKGGELLNMEEDYOXME	249
Dp	271	FLVALVAGVYETTTIMTFLAKFLTEPPLVALQLEKEHEKIRAKMSDSYSLEMSDYSKMP	330
QY	250	FTQCVISEALRCGNIVATVHRKATHDIKFEKYYLIPKGMKVPIPTAVYHLDPSLHENFEF	309
Dp	331	FTQCVAVETTELAVNIIGISVEFRRAATDVEIKGYKIPKGMKVFPSSRAVHLDPNHFKDARTE	390
QY	310	NPMNWTKT-----TAFGGSVRCPSGEGELGKLDIAFFLHLHLVLSYMW	350
Dp	391	NPMNMOSSNVTYTGPSNVETFPFGGSPRLCGPEGLARVALSVLELHRLVYGFQW	441

RESULT	2			
C90C_ARATH				
ID	C90C_ARATH	STANDARD;	PRT;	524 AA.
AC	G9M066; 023242;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Cytochrome P450 90C1 (EC 1.14.-.-) (ROTUNDIFOLI3).			
GN	ROT3 OR CYP90C1 OR At4G63680 OR C7A10.980 OR F2E13.220.			
OS	Arabidopsis thaliana (Mouse-ear cross).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.			
OX	NOBI_TaxId=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98361880; PubMed=9694802;			
RA	Kim G.-T., Tsukaya H., Uchimiya H.;			
RT	"The ROTUNDIFOLIA3 gene of Arabidopsis thaliana encodes a new member			
RT	of the cytochrome P-450 family that is required for the regulated			
RL	polar elongation of leaf cells."			
RL	Genes Dev. 12:2381-2391(1998).			
	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. COLUMBIA;			
RX	MEDLINE=20083488; PubMed=10617198;			
RA	Mayer K.F.X., Schneller C., Wandutt R., Murphy G., Volckaert G.,			
RA	Polh T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,			
RA	Harris B., Ansoerg W., Brandt P., Griwell L.A., Rieger M.,			
RA	Weitschalgartner M., de Simone Y., Obermaler B., Maché R., Mueller M.,			
RA	Kreis M., Delany M., Puigdomenech P., Watson M., Schmidheini T.,			
RA	Reichert B., Portetelle D., Perez-Alonso M., Bottry M., Bancroft I.,			
RA	Vos P., Hohenseil U., Zimmermann W., Wedler H., Ridley P.,			
RA	Langham S.-A., McCullagh B., Billam L., Robben J.,			
RA	Van der Schuren J., Gryonorez B., Chuang Y.-J., Vandenbusche F.,			
RA	Braeken M., Weltjens J., Voet M., Bastiaens I., Aert R., Defoor E.,			
RA	Wielensger T., Bothe G., Rampsberger U., Hilbert H., Braun M.,			
RA	Holzer E., Brandt A., Peters S., Van Staveren M., Dirse W.,			
RA	Moollman P., Klein Lankhorst R., Rose M., Hauf J., Koeltter P.,			
RA	Bernstein S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.			
RA	De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,			

RA Van Mouragu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., Mcalay K., Mayes R.,
RA Pettelt A., Rajandrem M.A., Lyne M., Benes V., Reichmann S.,
RA Bortova D., Bloeker H., Scharfe M., Grimm M., Loehmet T.-H.,
RA Dose S., de Haan M., Measre A.C., Schaefter M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Farmann B., Granderth K., Dauner D., Herzl A.,
RA Neumann S., Argilou A., Vitale D., Lignori R., Piravandi E.,
RA Massenet C., Ougley F., Glabaud G., Mendein A., Reiber R.,
RA Schenabl S., Hiller R., Schmidt W., Lecharny A., Abdour S.,
RA Cheldor F., Cooke R., Berger C., Montfort A., Casacubeta E.,
RA Gildons T., Weber N., Vandenbil M., Barques M., Terol J., Torres A.,
RA Perez-Perez A., Purrelle B., Bent E., Johnson S., Jacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
RA Frlman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la bastide M., Habermann K.,
RA Pannell L., Dedhia N., Gnoj L., Schut K., Huang E., Spiegel L.,
RA Sekon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Hamon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Mix P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Speth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antonou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
RA Granat S., Shohdy N., Hasegawa A., Hamed A., Rodhi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombie W.R.,
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RA thaliana.",
RL Nature 402:769-777(1999).
CC -I- FUNCTION: MIGHT BE INVOLVED IN THE BIOSYNTHESIS OF STEROIDS
CC IMPORTANT FOR THE POLAR ELONGATION OF CELLS DURING DEVELOPMENT.
CC -I- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum
CC (potential).
CC -I- TISSUE SPECIFICITY: UBIQUITOUS.
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB008097; BAA37167.1; -
DR EMBL; AL161589; CAB80304.1; ALT INT.
DR EMBL; Z99708; CAB18650.1; ALT INT.
DR EMBL; AL022141; CAA18139.1; ALT-SEQ.
DR InterPro; IPR001128; Cyt.P450.
DR Pfam; PF00067; P450.1.
DR PROSITE; PRO0385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KM Oxidoreductase; Monooxygenase; Transmembrane; Heme;
KW Endoplasmic reticulum; Multigene family;
TRANSMEM 4 24 POTENTIAL.
FT BINDING 463 463 HEME (BY SIMILARITY).
SEQUENCE 524 AA; 59389 MW; 550578908BDF272 CRC64;

Query Match	31.5%;	Score 638;	DB 1;	Length 524;
Best Local Similarity	32.2%;	Pred. No. 1.8e-36;		
Matches 141; Conservative	75;	Mismatches 156;	Indels 66;	Gaps 7;

[illegible]

[illegible]

```

RESULT      3
CP85_LYCES ID CP85_LYCES STANDARD: PRT: 464 AA.
AC Q43147;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Cytochrome P450 85 (EC 1.14.--) (Dwarf protein).
GN CYP85 OR D.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Assteridae; eustersids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV_CGR758;
RX MEDLINE=96266705; PubMed=8672892;
RA Bishop G.J., Harrison K., Jones J.G.D.;
RT "The tomato Dwarf gene isolated by heterologous transposon tagging
RT encodes the first member of a new cytochrome P450 family.";
RL Plant Cell 8:959-969(1996).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U54770; AA017070.1; -.
DR InterPro; IPR001128; Cyt_P450.
DR Pfam; PF00067; P450.1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450.1.
KM Oxidoreductase; monooxygenase; Membrane; Heme.
FT BINDING 414 HEME (BY SIMILARITY).
SQ SEQUENCE 464 AA; 53706 MW; D2B21AAAB7B14E94 CRC64;

```

Query Match	28.0%	Score 568.5;	DB 1;	Length 464;
Best Local Similarity	30.2%	Pred. No. 8.6e-32;		
Matches 130;	Conservative 78;	Mismatches 158;	Indels 65;	Gaps 9;
0y	1	MGMPIGETISFEKRRDSDIG-FTLQQRVSKYKGFVSNIGCGKAVVSCDELNNFIILQ	59	

```

Db 39 MGWPLFGTTEFLK-----LGPSFMKNQBARVGSFFKSHILGCPITVMSDSELNRIILV 92
Qy 60 NEGKLFSTDYKRAMHDILGKYSLLATGETIHRKLNVIISFINLTKSRPDFLHCENLST 119
Db 93 NEAKLVPGVQOSMIDILGKCNIAAVNGSAHKYKRGALLSTISPTMRDOLLPEKIDFMR 152
Qy 120 SILSKMKRCREVEPHKEKMPFLESVMNOLLSTIKREDPARLYVLODFLSYMKGISLPIR 179
Db 153 SHLNWMDN-KVIDIOEKNNKAPFLSLKQJLAGISTSLADFE-MSEFFNLVLGLSLPILN 210
Qy 180 LPTGTY-----TNAIKY-----RSNNRIHQNAIIEDMNN-----AIRREDFLDST 219
Db 211 LPTNTHHGFOARKIYIVLLTLEBRASKEIGHDLGYLMBEATRFKLTJDEMDILI 270
Qy 220 IS-----NEDEHNAIRAKKGDDGLLNMEDYQKMEFT 251
Db 271 ITIYSGEIVSTSMMAVKYLHDPVLELRKEIHMAIRKKKRPEDPIDYNDYRSMRFT 330
Qy 252 QCVISEALRCGNIKYTVRKATHDIKFREKEYIIPKGMVVPFTFAVHLDPSLHENPEFENP 311
Db 331 RAVILETSRLATIYNGVLARKTODMEINGIITIIPGMWIIYVTRBLANDPRLKYDPDYSFNP 350
Qy 312 MRWTKTT-----AFGGGVYRVCPEGDELKQJATFALHLVLVLSYRWKIKSDEMPANIHP 363
Db 391 WRMDKDSLSEHONSEFLVGGGTGRQCPGKELGVALISTELFHFTVTKRWBEIGIGDKLMKFR 450
Qy 364 VEPFRGMLEI 374
Db 451 VEAPNGRLIRV 461

```

RESULT	4
ID	C883_ARATH
AC	023051.1
DT	15-DEC-1998 (Rel. 37, Created)
DT	15-DEC-1998 (Rel. 37, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Cytochrome P450 88A3 (EC 1.14.-.-).
GN	CYP88A3 OR AT1G05160 OR YUP8H12.23.
OS	Arabidopsis thaliana (Mouse-ear cress).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC	eucotyledons II; Brassicales; Brassicaceae; Arabidopsis.
OX	NCBI_TaxID=3702;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=CV_COLUMBIA;
RX	MEDLINE=21016719; Pubmed=11130712;
RA	Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA	White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA	Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA	Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA	Dunn P., Etlz P., Feldlyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA	Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizler L.,
RA	Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA	Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,
RA	Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA	Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA	Miltschker J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA	Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA	Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA	Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA	Utechtack T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
RA	Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT	"Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT	thaliana.";
RL	Nature 408:816-820(2000).
CC	-----
CC	-I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC	the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AC000098; AAB71462.1; -
DR HSSP; P33006; 1CPT.
DR InterPro; IPR001128; CYL_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; P000385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR Oxidoreductase; Monooxygenase; Transmembrane; Heme.
FT TRANSMEM 6
FT BINDING 439 439 HEME (BY SIMILARITY).
SQ SEQUENCE 490 AA; 56409 MW; 7FD8CD7A8864D298 CRC64;

Query Match 21.3%; Score 432.5; DB 1; Length 490;
Best Local Similarity 24.9%; Pred. No. 1.7e-22;
Matches 106; Conservative 95; Mismatches 158; Indels 67; Gaps 10;

OY 1 MGPPIGETISFEKPHRSDISIGTFLQQRVSRGK--VEKSNICGKAVVSCDQELMFTL 58
DB 50 LGMPTIGMMLSLRAFKTSDDPSFTRLTKRYGPKGIYAHMFNPSTIIVTSDTCRRVL 109
OY 59 QNEGKLTSDYPRKAMHDILKYSLLATGELIRKLNKVIISFNLTKSKPDFLHCAENLS 118
DB 110 TDD-DAFKPGWFTSMELIGRKSFGISFEHKRLRLAAYNGHEALSTYIPYEAV 168
OY 119 ISILSKMKNCRVEFHEKYEKMTLSVWVNOQLSIKPEDPARLYLODELSTYKGFISLPI 178
DB 169 IYVLDKWKMGEEFLTHLRKLFRIIMYIFLSSSESNVMDA-LEREYALMYGVRAAMAV 227
OY 179 PLPGTYTNAIKVRNSRNHONAIIEDMNNAIREE--EGLDSIISNED----- 225
DB 228 NIPGFAIHRALAKKTLVAAFOSIVTERNOKONILSKKMDLNLVWKGDEGKTLD 287
OY 226 -----EHAAT--RAK-----KGDGL-LNME 243
DB 288 EEIIVLMLYNAGHSSGHTIMMATVFLQEHPEVLQRAKAQDEMILKSRPGKGLSLK 347
OY 244 DYKMEFQCVISEALRCGNIVKTHRKATHDIKREYIYIPGKVFPLFTVAHLDPSLH 303
DB 348 ETRKMEFLSÖVVDELRTYTFSLTAERAKTDVENMGYLIPPGMKVLTWFRVHIDPEVF 407
OY 304 ENPFENPMRW-----TKTTA---FGGAVVCPGGELGKLTQIAFLHHLVLSYRMKIKSD 355
DB 408 PDKRFDPARWNGFVPKAGAFLPFGAGSHLCPGNDLAKLEISIFLHFLKLYQVRSNP 467
OY 356 EMPIAH 361
DB 468 ECPVMY 473

RESULT 5
C881_MAIZE STANDARD; PRT; 519 AA.
AC 043246;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Cytochrome P450 88A1 (EC 1.14.-.-) (DWARF3 protein).
GN CYP88A1 OR D3.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV B73;
RX MEDLINE=96004534; PubMed=7549486;
RA Winkler R.G., Hellenfajris T.;

RT "The maize Dwarf3 gene encodes a cytochrome P450-mediated early step
RT in gibberellin biosynthesis."
RL Plant Cell 7:1307-1317(1995).
CC -1- PATHWAY: EARLY STEP IN GIBBERELLIN BIOSYNTHESIS.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ROOTS, DEVELOPING LEAVES, THE
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U32579; AAC49067.1; -
DR InterPro; IPR001128; CYL_P450.
DR Pfam; PF00067; P450; 1.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR Oxidoreductase; Monooxygenase; Transmembrane; Heme.
FT TRANSMEM 1
FT BINDING 466 466 HEME (BY SIMILARITY).
SQ SEQUENCE 519 AA; 57906 MW; 0E977A024316D95 CRC64;

Query Match 19.3%; Score 390.5; DB 1; Length 519;
Best Local Similarity 26.4%; Pred. No. 1.3e-19;
Matches 111; Conservative 69; Mismatches 160; Indels 81; Gaps 10;

OY 1 MGPPIGETISFEKPHRSDISIGTFLQQRVSRGK--VEKSNICGKAVV-----SCDQEL 53
DB 77 MGPPIVGMWMAFLRAFKSGKPDATFASVFRRGRTGVSFMFSSPTVLVTTAECKQVL 136
OY 54 NMFILQNEKLTSDYPRKAMHDILKYSLLATGELIRKLNKVIISFNLTKSKPDFLHC 113
DB 137 -----MDDDAFVVGPKPATVALVGRFSVAMPYDEHRRIRKLTLPAPINGFDALGYLPF 190
OY 114 AENLSISILKSW-KNCREVEFHEKYEKMTLSVWVNOQLSIKPEDPARLYLO-DELSYK 171
DB 191 IDRTYTSILRAADHGSSEFLELRKMTFKITIVQFLG--GADQATTPALRSTELMY 248
OY 172 GFTSLPIPLPGTYTNAIKVRNSRNHONAIIEDMNNAIREE-----DPLDSIISNED 224
DB 249 GMRAMAINDPGFAYRGALRAR--RLVAVLVQGVLDERRARARAGVSGGVDMDRILEAD 307
OY 225 -----EHAATIRAKKG 235
DB 308 EGRHLDDDELIVLVMYNAGHSSGHTIMMATVFLQENPDMFARAKAQEAIMKRSIPS 367
OY 236 DGEILNMDYOKMEFTQCVISEALRCGNIVKTHRKATHDIKREYIYIPGKVFPIPTA 295
DB 368 SQRGULTLRDFRKMEYLSÖVIDETLRLVNSIFSFRQATDVVNGVILPKGMKQVLMWYS 427
OY 296 VHLDSLHNPENPMRWTKTT-----AFGGAVVCPGGELGKLTQIAFLHHLVLSY 348
DB 428 VHMDDQVYVDPKFPSPRWEGSHSPRAGTFLARGLARLCPGNDLAKLEISVFLHFLGLY 487
OY 349 R 349
DB 488 K 488

RESULT 6
C120_SYNY3 STANDARD; PRT; 444 AA.
AC 059990;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative cytochrome P450 120 (EC 1.14.-.-).
GN CYP120 OR CYP OR STR0574.
OS Synechocystis sp. (strain PCC 6803).


```

OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-96127529; PubMed-8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 648 to 928 of the genome.";
RL DNA Res. 2:153-166(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE-97061201; PubMed-8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nieru K., Okumura S.,
RA Shino S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D64003; BAA10496.1;
DR InterPro: IPR001128; Cyt_P450.
DR Pfam: PF00067; P450.1; CYTOCHROME_P450.1.
DR PROSITE: PS00086; CYTOCHROME_P450.1.
DR Hypothetical protein; Oxidoreductase; Monooxygenase; Heme;
KW Complete proteome.
FT BINDING 391 HEME (BY SIMILARITY).
FT SEQUENCE 444 AA; 50578 MW; 8F62A0BED3B5ABDC CRC64;
SQ
Query Match 17.1%; Score 347.5; DB 1; Length 444;
Best Local Similarity 22.9%; Pred. No. 9e-11;
Matches 95; Conservative 76; Mismatches 164; Indels 79; Gaps 8;
OY 2 GMPGTFSTFFKPHRSDSITGLQORVSRKGVKFSNCGKAVVSCDQLNMFILONE 61
DB 19 GLPWLGETLNLND-----GDFGKRRQDFGPIETFLFGKNVFTISGALANRFLFTKE 72
OY 62 GLFTSDYPRKAMHDLGKYSLLATGEIHRKLNKVNIIISFINTKSPDPLCAENLSISI 121
DB 73 QETFOATWELSTRILLGRPALATOMGEIHRSRKILYQAF-LPRITDSTLPRMDGVQGY 131
OY 122 LKSMKNCREVEHKEVKMFTLSVWVNOQLSTIKREDPARLVLDLSTYKKGISLPIPLP 181
DB 132 LKONGKANEVVIYPLRRMTFVDAATLFEMGEKVSQNPQLFPM--FETVIOGLFSLPIPLP 189
OY 182 GGTGNATIVRSNRNINHOAIIEDMNNAIR-----EEDFLDLSIISNED----- 225
DB 190 MTLFGSKORARA-----LLALAEKTIKARQOOPSEEDALGILLARDNNOPLSLP 242
OY 226 -----EHAIRAKKGD-----GELLIMEDYQKME 249
DB 243 ELKDOIILLFAGHETLTSALSPCLLGGHSDIRERVAQEQNKLOLSELTAELTKMNP 302
OY 250 FPOCIYSELRCGNLYKTYHAKATHDIKFKEYVIRPKGVKVPFIFTAVHLDPSLHNPPEF 309
DB 303 YLDQVIOLEVLRLIPVGGGFRLLIDCDQFQGFHFPKGMVLSYQISOTNADPLDLPDEK 362
OY 310 NPMRMTKT-----TAFGGGVAVCPGEGELGLQIAFLHLHLVLSYRKKI 352

```

```

DB 363 DPERFPPDSATHNPPFAHVPGGGLRECLGKEFARLEMKLFATRLIQOEDMTL 416
RESULT 7
ID CP26_BRARE STANDARD; PRT; 492 AA.
AC P9739;
DT 15-DEC-1998 (rel. 37, Created)
DR 15-DEC-1998 (rel. 37, Last sequence update)
DT 15-DEC-1998 (rel. 37, Last annotation update)
DE Cytochrome P450 26 (EC 1.14.-.-) (Retinoic acid-metabolizing
DE cytochrome) (P450RA1) (Retinoic acid 4-hydroxylase).
GN CYP26.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-97094702; PubMed-8939936;
RA White J.A., Guo Y.-D., Baetz K., Beckett-Jones B., Bonasoro J.,
RA Hsu K.E., Dilworth F.J., Jones G., Petkovich M.;
RT "Identification of the retinoic acid-inducible all-trans-retinoic
RT acid 4-hydroxylase.";
RL J. Biol. Chem. 271:29922-29927(1996).
CC -1- FUNCTION: PLAYS A KEY ROLE IN RETINOIC ACID METABOLISM. ACTS ON
CC RETINOLIDS, INCLUDING ALL-TRANS-RETINOIC ACID (RA) AND ITS
CC STEREOISOMER 9-CIS-RA. CAPABLE OF 4-HYDROXYLATION, HYDROXYLATION,
CC RESPONSIBLE FOR GENERATION OF SEVERAL HYDROXYLATED FORMS OF RA,
CC INCLUDING 4-OH-RA AND 4-OXO-RA.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- INDUCTION: BY RETINOIC ACID.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U68234; AAC60045.1;
DR ZFIN: ZDB-GENE-990415-44; CYP26.
DR InterPro: IPR001128; Cyt_P450.
DR Pfam: PF00067; P450.1;
DR PRINTS: PR00385; P450.1.
DR PROSITE: PS00086; CYTOCHROME_P450.1.
DR Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome;
KW Endoplasmic reticulum.
FT BINDING 438 HEME (BY SIMILARITY).
FT SEQUENCE 492 AA; 56281 MW; FD471435B2F30509 CRC64;
SQ
Query Match 14.5%; Score 294.5; DB 1; Length 492;
Best Local Similarity 24.3%; Pred. No. 4.1e-13;
Matches 111; Conservative 59; Mismatches 190; Indels 97; Gaps 13;
OY 1 MGPPIGTFISFFKPHRSDSITGLQORVSRKGVKFSNCGKAVVSCDQLNMFILONE 60
DB 49 MGLPFGTGLQILDRK-----FLRMKRQKYGCIYKTLFGFNPRVWVAGADNVRQILLG 103
OY 61 ECKLFTSDYPRKAMHDLGKYSLLATGEIHRKLNKVNIIISFINTKSPDPLCAENLSISI 120
DB 104 EHKLVSVQMPASVRLSDTSLNVHGVQHKKKKAAIMAFS-RDALEHYIPVIOQEVKS 162
OY 121 LKSN---KNCREVEHKEVKMFTLSVWVNOQLSTIKPE---DPAFLVLDLSTYKKG 173
DB 163 AIOEWLQKDC--VAVYEMKMKLMFRIMRLIGREPQIKTDEGL--VEAFEEIMINL 218
OY 174 ISLPLPLGCTGTNAIKRS-----NRNIHONAIIEDMNNAIREDPDLDSIIS--ED 224

```

```
DB 219 FSLPIDVPSGLYRGRLARNFHLSKIEENIRKKIODDNNENQKYKDALQLLIENSRSD 278
OY 225 EENHAIRAKKGDEL-----LNMDYOK-----247
DB 279 EEPFLQAMEKATELLEFGHEHTTASTSLVMEFLGINTFVQKVRREEQEKVMGYTPG 338
OY 248 -----MEFTQVISEALRCGNIVKTVHRKATHDIEKFEYIPKGMKVEPIETAVH 297
DB 339 KGLSMELDLQAKYTCVIVETLRINPVPGGFRVAKTFELNGYQIPKGMNVIYSICDH 398
OY 298 LDPSLHENPEFENPMKWT-----TAAGGVAVCPGGELGKQIAFFLHLVLS 347
DB 399 DVADVPFNKEEFQPERFERMGKLEDGSRFNYIPFGGSRMCVCKEFAKVLKTLFVELTQH 458
OY 348 YRRKIKSDEN-----PIAHYVEFKRGMLEIEPTKE 379
DB 459 CMMILSNPPMTKGTPTTP-----VDNLPTKE 486

RESULT 8
CP26_MOUSE STANDARD: PRT: 497 AA.
AC 053127; 09R1F4;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome P450 26 (EC 1.14.-.-) (Retinoic acid-metabolizing
  cytochrome) (P450RA1) (Retinoic acid 4-hydroxylase).
GN CYP26A1 OR CYP26 OR P450RA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H;
RX MEDLINE=97392446; PubMed=9250660;
RA Fujii H., Sato T., Kaneko S., Gotoh O., Fujii-Kuriyama Y., Osawa K.,
RA Kato S., Hamada H.;
RT "Metabolic inactivation of retinoic acid by a novel P450
  differentially expressed in developing mouse embryos.";
RL EMBO J. 16:4163-4173(1997).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98113212; PubMed=9442090;
RA Abu-Abed S.S., Beckett B.R., Chiba H., Chithalen J.V., Jones G.,
RA Metzger D., Chambon P., Petkovich M.;
RT "Mouse P450RA1 (CYP26) expression and retinoic acid-inducible retinoic
  acid metabolism in F9 cells are regulated by retinoic acid receptor
  gamma and retinoid X receptor alpha.";
RL J. Biol. Chem. 273:2409-2415(1998).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Tooth;
RX MEDLINE=20515038; PubMed=11063033;
RA Paine C.T., Paine M.L., Snead M.L.;
RT "Identification of tufelin- and amelogenin-interacting proteins using
  the yeast two-hybrid system.";
RL Connect. Tissue Res. 38:257-267(1998).
[4]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (Aug-2001) to the EMBL/Genbank/DDay databases.
CC -!- FUNCTION: Plays a key role in retinoic acid metabolism. Acts on
  retinoids, including all-trans-retinoic acid (RA) and its
  stereoisomer 9-cis-RA. Capable of both 4-hydroxylation and 18-
  hydroxylation. Responsible for generation of several hydroxylated
  forms of RA, including 4-OH-RA, 4-oxo-RA, AND 18-OH-RA.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: BY RETINOIC ACID.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
-----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y12657; CAA73206.1; -.
DR EMBL; AF115769; AA017217.1; -.
DR EMBL; BC012673; AA012673.1; -.
DR MGD; MGI:1096359; Cyp26.
DR InterPro; IPR001128; CYP_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome;
KW Endoplasmic reticulum.
FT BINDING 442 442 HEME (POTENTIAL).
FT CONFLICT 9 9 S -> T (IN REF. 3).
FT CONFLICT 154 154 L -> P (IN REF. 4).
FT CONFLICT 356 356 I -> T (IN REF. 4).
FT CONFLICT 492 492 Y -> H (IN REF. 4).
SQ SEQUENCE 497 AA; 56177 MW; 33B07D/C29134471 CRC64;

Query Match 14.4%; Score 292.5; DB 1; Length 497;
Best Local Similarity 23.2%; Pred. No. 5.7e-13;
Matches 106; Conservative 61; Mismatches 187; Indels 103; Gaps 10;

OY 1 MGNPFIGETISFEKPHRSDSIGTFLQQRVSRGYKFSNICGKAVYSCDQELNMFILON 60
DB 49 MGPFPGETLQWVLRK-----FLQMKRRKYGFYKTHLFGFPTVAVGADNVRILLG 103
OY 61 ECKLFTSDPKAMHDLGKXSLLATGELHRRKKNVIFINITSKPRFLCAENLTS 120
DB 104 EHRLVSVHWPASVYRTLLGAGCLSNLHDSHKQKKVIMQAFSEALQCYLVIAEVS-S 162
OY 121 ILKSMKNC--REVEFHKEMKFTLSVYNQLSIKP-----EDPARLYVLDPLSYMG 172
DB 163 CLEQWLSGCEGRLVYPEVKRLMFRIMRLILGCEGPGAGGDEQL--VEAFEEMTN 220
OY 173 FLSPLPLPGGTYNAIKYRSNRNIHONAIIEDMNAIRE-----EDFLDSIISN 222
DB 221 LFLSLPIDVPSGLYRGVKA---NLIHRIEENIRAKIRRLQATEPDDGCKALQLLIRH 277
OY 223 -----EDEHNAIRA 232
DB 278 SWERGERLDMQALKOSSTELLEFGHEHTTASATSLITYIGLPHVYLQKREIKSKGLIC 337
OY 233 KKGDGELLNMEYQKMEFTQCVISEALRCGNIVKTVHRKATHDIEKFEYIPKGMKVEPI 292
DB 338 KSNQDKKLDMELEQKLTIGCVIKETLRINPVPGGFRVAKTFELNGYQIPKGMNVIYS 397
OY 293 FTAVHLDPSSLHENPEFENPMKWT-----KTAFSGGVAVCPGGELGKQIAFFLHL 342
DB 398 ICDTHVAADIFTNKEEFNDRFVPHPEDASRSEPTFGGSLRSCGKKEFAKTLKTLFTV 457
OY 343 HVLVSRWKIKS-----DEMPIAPHYVE 365
DB 458 ELARHCDWOLLNGPPTMKTSPTVYVDNLPARFTYQ 494

RESULT 9
CP26_XENLA STANDARD: PRT: 492 AA.
AC 093323;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cytochrome P450 26 (EC 1.14.-.-) (Retinoic acid degrading enzyme
  CYP26) (XCP26) (Retinoic acid converting enzyme) (RACE).
GN CYP26A1 OR CYP26.
-----
```

OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodidae; Xenopus.
NCBI_TaxID=8355;
RN
RP
RX MEDLINE=99077803; PubMed=9857192;
RA Hollemann T., Chen Y., Grunz H., Piefer T.;
RT "Regionalized metabolic activity establishes boundaries of retinolic acid signaling";
RL EMO J. 17:7361-7372(1998).
CC
CC -1- FUNCTION: Plays a key role in retinolic acid metabolism (By similarity).
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum (By similarity).
CC -1- INDUCTION: By retinolic acid.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
DR EMBL: AF057566; AAC25158.1; -
DR InterPro: IPR001128; Cyt_P450.
DR Pfam: PF00067; P450.1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450.1.
KW Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome; Endoplasmic reticulum.
FT BINDING 437 437 HEME (POTENTIAL).
FT SEQUENCE 492 AA; 55459 MW; D1D4BB7651BF2D3E CRC64;
Query Match 13.8%; Score 280; DB 1; Length 492;
Best Local Similarity 24.1%; Pred. No. 4e-12;
Matches 104; Conservative 54; Mismatches 176; Indels 98; Gaps 12;
OY 1 MGPPIGETISFFKRRSDISCTFLOQVRVSRGKVFKNICGKAVSCDQELNFIION 60
DB 49 MGLPFGFGLWAVLQRR-----RFLQVRSQYGRITVTHLFGSPVRYTGAENVQILMG 103
OY 61 BECKLTSDYPRKMHDIQKYSLLATGEIHRKKNVYI-----SPINFKSKPFLICA 114
DB 104 EHKLVSVHMPASVKTILLGCLSNLHNDHKYTKVIAOAFSRBALANYVPMEEVRCS 163
OY 115 ENLSISILSKMKN--CREVEFHKVEKMTLSVMNOLSLINPEDPARL-----YVLQD 165
DB 164 VNL-----WLGSGPC--VLVYPAIKRMFMRIAMRLLLGC--DPQRNDRQEEFLLEA 211
OY 166 FLSTYKGFSLPIPLPGTYNAIVR-----
DB 212 PEMSRNLFSLPIDVPSGLRGLARNLHAQIEENIKELQRPDEHCKDALQLLIDY 271
OY 193 SNRN--IHONAIIDMNAI-----REDFLDSIISNEDEHNAI 230
DB 272 SRNKEPILNQLAKSATALLFGHGHTASATSLTSLFALHKDVLK--VRKELETGCL 330
OY 231 RAKKGDGELLMEDYOKMEFTQCVISEALRCGNIVKTVHRKATHDIKFEKVIYPRGMVYF 290
DB 331 STKPEKKEKLSIEVLQQLKTSYCVIKETRLRLSPVAGGFRVALKTFVLNGYDIPKGMVY 390
OY 291 PFTFVHLDPSLHNPFEENPKMTKT-----TAFGGVAVRCPOGELGKQIAFF 340
DB 391 YSIADTHGADLPDTRKPNDRFLPLPRDSSRGGFIPFGGGVRCIGKEPAKILKVF 450
OY 341 LHHLVLSYRKK 352
DB 451 VVELRCNCDMEL 462

RESULT 10
C26A_HUMAN
ID C26A_HUMAN STANDARD: PRT: 512 AA.
AC Q9NR63; Q9NP41; -
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cytochrome P450 26A2 (EC 1.14.-.-) (P450RA1-2) (Retinolic-acid metabolizing cytochrome).
DE CYP26A2 OR P450RA12.
GN Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN
RP
RX MEDLINE=20300913; PubMed=10823918;
RA White J.A., Ramshaw H., Taimi M., Stangie W., Zhang A., Everingham S., Creighton S., Tam S.-P., Jones G., Petkovich M.;
RT "Identification of the human cytochrome P450, P450RA1-2, which is predominantly expressed in the adult cerebellum and is responsible for all-trans-retinolic acid metabolism";
RL Proc. Natl. Acad. Sci. U.S.A. 97:6403-6408(2000).
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=20300913; PubMed=10823918;
RA White J.A., Ramshaw H., Taimi M., Stangie W., Zhang A., Everingham S., Creighton S., Tam S.-P., Jones G., Petkovich M.;
RT "Identification of the human cytochrome P450, P450RA1-2, which is predominantly expressed in the adult cerebellum and is responsible for all-trans-retinolic acid metabolism";
RL Proc. Natl. Acad. Sci. U.S.A. 97:6403-6408(2000).
RN
RP SEQUENCE FROM N.A.
RA Bourne S., Bauer C., Pape K., Jones T.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Plays a key role in retinolic acid metabolism. Involved in the specific inactivation of all-trans-RA. Responsible for generation of several hydroxylated forms of RA, including 4-OH-RA, 4-oxo-RA, and 18-OH-RA.
CC -1- ENZYME REGULATION: Has a preferred activity towards the following substrates: all-trans-RA > 9-cis-RA > 13-cis-RA.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum (By similarity).
CC -1- TISSUE SPECIFICITY: Highly expressed in brain, particularly in the cerebellum and pons.
CC -1- INDUCTION: By retinolic acids (RA).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
DR EMBL: AF252297; AAF76003.1; -
DR EMBL: AC007002; AAF65576.1; -
DR MIM: 605207; -
DR InterPro: IPR001128; Cyt_P450.
DR Pfam: PF00067; P450.1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450.1.
KW Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome; Endoplasmic reticulum.
FT BINDING 441 441 HEME (POTENTIAL).
FT CONFLICT 68 68 Q -> QKCTLRVWMLQ (IN REF. 2).
FT SEQUENCE 512 AA; 57512 MW; A0DD1D9944E6726F CRC64;
Query Match 13.7%; Score 278.5; DB 1; Length 512;
Best Local Similarity 23.0%; Pred. No. 5.3e-12;
Matches 105; Conservative 69; Mismatches 161; Indels 121; Gaps 15;
OY 1 MGPPIGETISFFKRRSDISCTFLOQVRVSRGKVFKNICGKAVSCDQELNFIION 60
DB 54 MGLPFGFGLWAVLQRR-----RFLQVRSQYGRITVTHLFGSPVRYTGAENVQILMG 108

NCBI_TaxId=9606;
 [1]
 RX SEQUENCE FROM N.A. MEDLINE=97373542; PubMed=9228017;
 RX White J.A., Beckett-Jones B., Guo Y.-D., Dilworth F.J., Bonasoro J.,
 RA Jones G., Petkovich M.;
 RT "cDNA cloning of human retinoic acid-metabolizing enzyme (hP450RA1)
 JT identifies a novel family of cytochromes P450.";
 RL J. Biol. Chem. 272:18538-18541(1997).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=98380037; PubMed=9716180;
 RX Sonnenfeld E., van der Burg C.E., van der Leede B.M., Schuijkes R.K.,
 RA Petkovich M., van der Burg B., van der Saag P.T.;
 RT "human retinoic acid (RA) 4-hydroxylase (CYP26) is highly specific for
 RT all-trans-RA and can be induced through RA receptors in human breast
 RT and colon carcinoma cells.";
 RL Cell Growth Differ. 9:629-637(1998).
 RN [3]
 RN TISSUE SPECIFICITY.
 RP MEDLINE=99045433; PubMed=9826557;
 RX Trofinova-Grifflin M.E., Juchau M.R.;
 RA "Expression of cytochrome P450RA1 (CYP26) in human fetal hepatic and
 RT cephalic tissues.";
 RL Biochem. Biophys. Res. Commun. 252:487-491(1998).
 CC -1- FUNCTION: Plays a key role in retinoic acid metabolism. Acts on
 CC retinoids, including all-trans-retinoic acid (RA) and its
 CC stereoisomer 9-cis-RA. Capable for generation of several hydroxylated
 CC forms of RA, including 4-OH-RA, 4-oxo-RA, AND 18-OH-RA.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -1- TISSUE SPECIFICITY: Highest levels in adult liver, heart,
 CC pituitary gland, adrenal gland, placenta and regions of the brain.
 CC -1- INDUCTION: By retinoic acid.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).
 CC Or send an email to license@sib-sib.ch.
 CC -----
 CC EMBL: AF005418; AAB88881.1; -
 DR HSSP: P33006; LCPY.
 DR MIM: 602239; -
 DR InterPro: IPR001128; CYP_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome;
 KW Endoplasmic reticulum
 FT BINDING 442 442 HEME (POTENTIAL).
 FT SEQUENCE 497 AA: 56162 MW: 586084B24B25A93 CRC64:

[illegible]

```

RESULT 13
CPA3_RAT
ID CPA3_RAT STANDARD: PRT: 494 AA.
AC P20812.
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Cytochrome P450 2A3 (EC 1.14.14.1) (CYP1A3) (Coumarin 7-hydroxylase).
GN CPA3 OR CYP2A-3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Lung;
RX MEDLINE=90356430; PubMed=2388852;
RA Ueno T., Gonzalez F.;
RT "Complete sequence of the rat CYP2A3 gene specifically transcribed in
RT lung."
RL Nucleic Acids Res. 18:4623-4624(1990).
RN [2]
RP SEQUENCE OF 8-494 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=89323084; PubMed=2751996;
RA Kimura S., Kozak C.A., Gonzalez F.J.;
RT "Identification of a novel P450 expressed in rat lung: cDNA cloning
RT and sequence, chromosome mapping, and induction by
RT 3-methylcholanthrene."
RL Biochemistry 28:3798-3803(1989).
CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATY
CC ACIDS, AND XENOBIOTICS.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) -> ROH +
CC oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- TISSUE SPECIFICITY: LUNG.
CC -1- INDUCTION: BY 3-METHYLCOLANTHRENE (3MC).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL: M33190; AAA41022.1; -.
CC DR EMBL: J02852; AAA88511.1; -.

```

DR PIR: A32030; A32030.
 DR PIR: S15056; S15056.
 DR HSSP: P00179; 1DT6.
 DR InterPro: IPR001128; Cyt_P450.
 DR Pfam: PF00067; P450.1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KM Microsome; Endoplasmic reticulum.
 FT BINDING 439 439 HEME.
 FT CONFLICT 109 109 M -> G (IN REF. 2).
 FT CONFLICT 335 335 I -> N (IN REF. 2).
 SQ SEQUENCE 494 AA; 56510 MW; F71855CBB602672F CRC64;

Query Match 11.4%; Score 230.5; DB 1; Length 494;
 Best Local Similarity 22.4%; Pred. No. 9,4e-09;
 Matches 102; Conservative 67; Mismatches 147; Indels 139; Gaps 14;

QY 4 PRIGETISFFKPHRSDSIFTLQQRVSRYGKFKSNICGKAVSCDQLNMFILQNEGK 63
 DB 41 PFIENYLIQL-----NTEKWSLSIMKISQRYGPVFTIHLGPRRVVVLGGQAVKALVDQAE 96
 QY 64 LFTSDYKAMHDILGK-YSLLATGELHKKLNVIISFINTLTKSKPPLHCAENLSTISIL 122
 DB 97 EFSGRGEQATFDMWFKGYGVAFSSGERAKQLR-----RFSIATL 135
 QY 123 KSW---KNCREVEFHKEV-----KMFTLSVMVNOQLSINKPEDPARLYVL 163
 DB 136 RDEFGVGRGIEERIOEAGLIESFKTNGALDPTLYLSRTVSNVSSIVFGDRPED 195
 QY 164 QDLSTYK-----GFTSLPPLPG-----TGYNIAIKVRNRM 196
 DB 196 KEFLSLIRMMIGSFQFATSTGOLYEMFSSVMKHLPGQQAPELQGLDEFL---TKK 251
 QY 197 IHNVAILEDNNNAIREDEFDSISNDEE----- 226
 DB 252 VEONQRIIDPNS---PDDFIDSLIRMLERKKNPTEFYKKNLVLTLLNLFAGTETVST 308
 QY 227 -----HAIRAKK-----GDGELLNWDYQKMEFTQCIYSPALRCGNIVKT 267
 DB 309 TLKRGFLLMKHPDIEAKVHEIDRVIGRNQAKYEDRMKMPTEAVITHETQRPADHIRM 368
 QY 268 -VHRKATHDKKEKYVTPKWKVPRPTAVHLDPDSLHENPFPNPM-----RMTKTGA 319
 DB 369 GLARVTKDKFRREFLLPKGTVEVPMGLKDPKFFSNPNDENPKHFLDDKQGFKKSDA 428
 QY 320 ---FGGGRVRCPPGELGKLIQIAFFLHHLVLSYRMK 351
 DB 429 FVFPSTGKRCYCPGEGGLARMLFLFLTNIMQNFQCFK 463

RESULT 14
 ID CPAT_HUMAN STANDARD; PRT; 494 AA.
 AC P20853; O13121.
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Cytochrome P450 2A7 (EC 1.14.14.1) (CYP1A7) (P450-IIIA).
 GN CYP2A7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=90212623; PubMed=2322567;
 RA Yamano S., Tatsuno J., Gonzalez F.J.;
 RT "The CYP2A7 gene product catalyzes coumarin 7-hydroxylation in human
 liver microsomes.";
 RL Biochemistry 29:1322-1329(1990).

RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95397851; PubMed=7668294;
 RA Fernandez-Salguero P., Hoffman S.M., Cholesterol S., Mohrenweiser H.,
 RA Raunio H., Rautio A., Pelkonen O., Huang J.D., Evans W.E., Idle J.R.;
 RT "A genetic polymorphism in coumarin 7-hydroxylation: sequence of the
 RT human CYP2A7 genes and identification of variant CYP2A6 alleles.";
 RL Am. J. Hum. Genet. 57:651-660(1995).
 CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
 CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
 CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
 CC OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
 CC ACIDS, AND XENOBIOTICS.
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -1- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER AND OTHER
 CC TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS, PESTICIDES,
 CC AND CARCINOGENS.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).
 CC -----
 DR EMBL: M33317; AAA52138.1; -;
 DR EMBL: U22029; AAB40520.1; -;
 DR PIR: C34271; C34271.
 DR HSSP: P00179; 1DT6.
 DR MTM: 123960; -;
 DR InterPro: IPR001128; Cyt_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KM Microsome; Endoplasmic reticulum.
 FT BINDING 439 439 HEME (BY SIMILARITY).
 FT CONFLICT 108 108 D -> E (IN REF. 2).
 FT CONFLICT 163 163 S -> T (IN REF. 2).
 FT CONFLICT 226 226 V -> L (IN REF. 2).
 FT CONFLICT 242 242 Q -> L (IN REF. 2).
 FT CONFLICT 402 403 PS -> LR (IN REF. 2).
 FT CONFLICT 409 409 Q -> R (IN REF. 2).
 FT CONFLICT 418 419 DD -> GE (IN REF. 2).
 FT CONFLICT 426 426 S -> R (IN REF. 2).
 FT CONFLICT 435 435 G -> R (IN REF. 2).
 FT CONFLICT 438 438 Y -> N (IN REF. 2).
 FT CONFLICT 475 475 S -> P (IN REF. 2).
 FT CONFLICT 479 479 G -> V (IN REF. 2).
 SQ SEQUENCE 494 AA; 56408 MW; 4094386AA07F9016 CRC64;

Query Match 11.4%; Score 230.5; DB 1; Length 494;
 Best Local Similarity 23.7%; Pred. No. 9,4e-09;
 Matches 108; Conservative 62; Mismatches 157; Indels 129; Gaps 19;

QY 4 PRIGETISFFKPHRSDSIFTLQQRVSRYGKFKSNICGKAVSCDQLNMFILQNEGK 63
 DB 41 PFIENYLIQL-----NTEKWSLSIMKISQRYGPVFTIHLGPRRVVVLGGQAVKALVDQAE 96
 QY 64 LFTSDYKAMHDILGK-YSLLATGELHKKLNVIISFINTLTKSKPPLHCAENLSTISIL 122
 DB 97 EFSGRGEQATFDMWFKGYGVAFSSGERAKQLR-----RFSIATL 135
 QY 101 INTTSK-----PDL--HCAENLSTILSKMKNCREVEFHKEVME--TISVMNQLL 150
 DB 157 IAIKRSKGANIDPTFFLSRTVSNVSSIVFGDR-----FYDEDFSLSLSMIG-IF 209
 QY 151 SKRPEDPARLYVLQDLSTYKGFISLPL-----LPQGTYNIAIKVRSN-RNIHONAI 202

```

Db 210 QFTSTGQLEYEM--FSSVAK---HLPGPOQAQFKLLQGLEDFIAKVEHNQRTLDNS- 263
QY 203 IEDMNNAIREDDELDSIISNEDEE----- 226
Db 264 -----PQDITDSFLIMQEEKNPNTPEFLKINLMSTLNLFTAGTETVSTTLRYGF 314
QY 227 -----HAATIAAK-----GGCELLNMDYOKMEFTQCVISALRCGNIV-KTVHRA 272
Db 315 LLMKHEVEAKVHEEDIRYIGKNRQPKFEDRTKMPYMEAVIHIEIQFGDVIPMSLARV 374
QY 273 THDIKPEYIIPKGMKYEPIFTAVHIDPSLHNPFEENPM-----RMKRTA---EGG 322
Db 375 KKTPEKDFELPKGTVEFPMGLSVLRDPSFSPNPQDPHFLDQKQFKKSDAFVFPFSI 434
QY 323 GVRVCPGSELGLKLIQIAFFLHHLVSYRMKIKSDMP 358
Db 435 GKRYCFCGEGCLAMREL--FLFTTYVMQNFRLKSSQSP 468

RESULT 15
CP44_MOUSE
ID CP44_MOUSE STANDARD: PRT: 494 AA.
AC PLS392:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytochrome P450 2A4 (EC 1.14.14.1) (CYP2A4) (Testosterone 15-alpha-
GN hydroxylase) (P450-15-alpha) (P450-11A3.1).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89197954; PubMed=2703500;
RA Lindberg R., Burkhardt B., Ichikawa T., Negishi M.;
RT "The structure and characterization of type I P-450(15) alpha gene as
RT major steroid 15 alpha-hydroxylase and its comparison with type II P-
RT 450(15) alpha gene."
RL J. Biol. Chem. 264:6465-6471(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88153731; PubMed=3346244;
RA Squires E.J., Negishi M.;
RT "Reciprocal regulation of sex-dependent expression of testosterone 15
RT alpha-hydroxylase (P-450(15) alpha) in liver and kidney of male mice
RT by androgen. Evidence for a single gene."
RL J. Biol. Chem. 263:4166-4171(1988).
RN [3]
RP MUTAGENESIS.
RX MEDLINE=89281737; PubMed=2733794;
RA Lindberg R., Negishi M.;
RT "Alteration of mouse cytochrome P450och substrate specificity by
RT mutation of a single amino-acid residue."
RL Nature 339:632-634(1989).
RN [4]
RP TISSUE SPECIFICITY.
RX MEDLINE=99421934; PubMed=10490589;
RA Lavery D.J., Loper-Molina L., Margueron R., Fleury-Olela F.,
RA Conquet F., Schibler U., Bonfils C.;
RT "Circadian expression of the steroid 15 alpha-hydroxylase (Cyp2a4) and
RT coumarin 7-hydroxylase (Cyp2a5) genes in mouse liver is regulated by
RT the PAR leucine zipper transcription factor DBP."
RL Mol. Cell. Biol. 19:6488-6499(1999).
CC -1- FUNCTION: HIGHLY ACTIVE IN THE 15-ALPHA-HYDROXYLATION OF
CC TESTOSTERONE. ALSO ACTIVE IN THE 15-ALPHA-HYDROXYLATION OF
CC PROGESTERONE AND ANDROSTENEDIONE. LITTLE OR NO ACTIVITY ON
CC CORTICOSTERONE, PREGNENOLONE, DEHYDROPIANDROSTERONE, ESTRADIOL OR
CC ESTRIDOL.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +

```

```

CC oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- TISSUE SPECIFICITY: KIDNEY AND LUNG. EXPRESSED IN LIVER, WITH A
CC STRONG CIRCADIAN RHYTHMICITY. CIRCADIAN EXPRESSION IS REGULATED BY
CC DBP.
CC -1- MISCELLANEOUS: THERE ARE ONLY 11 DIFFERENCES BETWEEN THE SEQUENCE
CC OF TESTOSTERONE 15-ALPHA-HYDROXYLASE AND THAT OF COUMARIN 7-
CC HYDROXYLASE. BY SITE-DIRECTED MUTAGENESIS IT HAS BEEN SHOWN THAT
CC MODIFICATION OF POSITION 209 IS SUFFICIENT TO CONVERT THE
CC SPECIFICITY OF THE TWO FORMS OF THE ENZYME.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M26208; AAA37797.1; -.
DR EMBL: M25146; AAA37797.1; JOINED.
DR EMBL: M25147; AAA37797.1; JOINED.
DR EMBL: M26202; AAA37797.1; JOINED.
DR EMBL: M26203; AAA37797.1; JOINED.
DR EMBL: M26205; AAA37797.1; JOINED.
DR EMBL: M26206; AAA37797.1; JOINED.
DR EMBL: M26207; AAA37797.1; JOINED.
DR EMBL: M26209; AAA37797.1; JOINED.
DR EMBL: M26204; AAA37797.1; JOINED.
DR EMBL: M26201; AAA37797.1; JOINED.
DR EMBL: M26200; AAA37797.1; JOINED.
DR EMBL: M26202; AAA37797.1; JOINED.
DR EMBL: M26203; AAA37797.1; JOINED.
DR EMBL: M26204; AAA37797.1; JOINED.
DR EMBL: M26205; AAA37797.1; JOINED.
DR EMBL: M26206; AAA37797.1; JOINED.
DR EMBL: M26207; AAA37797.1; JOINED.
DR EMBL: M26208; AAA37797.1; JOINED.
DR EMBL: M26209; AAA37797.1; JOINED.
DR EMBL: M26210; AAA37797.1; JOINED.
DR EMBL: M26211; AAA37797.1; JOINED.
DR EMBL: M26212; AAA37797.1; JOINED.
DR EMBL: M26213; AAA37797.1; JOINED.
DR EMBL: M26214; AAA37797.1; JOINED.
DR EMBL: M26215; AAA37797.1; JOINED.
DR EMBL: M26216; AAA37797.1; JOINED.
DR EMBL: M26217; AAA37797.1; JOINED.
DR EMBL: M26218; AAA37797.1; JOINED.
DR EMBL: M26219; AAA37797.1; JOINED.
DR EMBL: M26220; AAA37797.1; JOINED.
DR EMBL: M26221; AAA37797.1; JOINED.
DR EMBL: M26222; AAA37797.1; JOINED.
DR EMBL: M26223; AAA37797.1; JOINED.
DR EMBL: M26224; AAA37797.1; JOINED.
DR EMBL: M26225; AAA37797.1; JOINED.
DR EMBL: M26226; AAA37797.1; JOINED.
DR EMBL: M26227; AAA37797.1; JOINED.
DR EMBL: M26228; AAA37797.1; JOINED.
DR EMBL: M26229; AAA37797.1; JOINED.
DR EMBL: M26230; AAA37797.1; JOINED.
DR EMBL: M26231; AAA37797.1; JOINED.
DR EMBL: M26232; AAA37797.1; JOINED.
DR EMBL: M26233; AAA37797.1; JOINED.
DR EMBL: M26234; AAA37797.1; JOINED.
DR EMBL: M26235; AAA37797.1; JOINED.
DR EMBL: M26236; AAA37797.1; JOINED.
DR EMBL: M26237; AAA37797.1; JOINED.
DR EMBL: M26238; AAA37797.1; JOINED.
DR EMBL: M26239; AAA37797.1; JOINED.
DR EMBL: M26240; AAA37797.1; JOINED.
DR EMBL: M26241; AAA37797.1; JOINED.
DR EMBL: M26242; AAA37797.1; JOINED.
DR EMBL: M26243; AAA37797.1; JOINED.
DR EMBL: M26244; AAA37797.1; JOINED.
DR EMBL: M26245; AAA37797.1; JOINED.
DR EMBL: M26246; AAA37797.1; JOINED.
DR EMBL: M26247; AAA37797.1; JOINED.
DR EMBL: M26248; AAA37797.1; JOINED.
DR EMBL: M26249; AAA37797.1; JOINED.
DR EMBL: M26250; AAA37797.1; JOINED.
DR EMBL: M26251; AAA37797.1; JOINED.
DR EMBL: M26252; AAA37797.1; JOINED.
DR EMBL: M26253; AAA37797.1; JOINED.
DR EMBL: M26254; AAA37797.1; JOINED.
DR EMBL: M26255; AAA37797.1; JOINED.
DR EMBL: M26256; AAA37797.1; JOINED.
DR EMBL: M26257; AAA37797.1; JOINED.
DR EMBL: M26258; AAA37797.1; JOINED.
DR EMBL: M26259; AAA37797.1; JOINED.
DR EMBL: M26260; AAA37797.1; JOINED.
DR EMBL: M26261; AAA37797.1; JOINED.
DR EMBL: M26262; AAA37797.1; JOINED.
DR EMBL: M26263; AAA37797.1; JOINED.
DR EMBL: M26264; AAA37797.1; JOINED.
DR EMBL: M26265; AAA37797.1; JOINED.
DR EMBL: M26266; AAA37797.1; JOINED.
DR EMBL: M26267; AAA37797.1; JOINED.
DR EMBL: M26268; AAA37797.1; JOINED.
DR EMBL: M26269; AAA37797.1; JOINED.
DR EMBL: M26270; AAA37797.1; JOINED.
DR EMBL: M26271; AAA37797.1; JOINED.
DR EMBL: M26272; AAA37797.1; JOINED.
DR EMBL: M26273; AAA37797.1; JOINED.
DR EMBL: M26274; AAA37797.1; JOINED.
DR EMBL: M26275; AAA37797.1; JOINED.
DR EMBL: M26276; AAA37797.1; JOINED.
DR EMBL: M26277; AAA37797.1; JOINED.
DR EMBL: M26278; AAA37797.1; JOINED.
DR EMBL: M26279; AAA37797.1; JOINED.
DR EMBL: M26280; AAA37797.1; JOINED.
DR EMBL: M26281; AAA37797.1; JOINED.
DR EMBL: M26282; AAA37797.1; JOINED.
DR EMBL: M26283; AAA37797.1; JOINED.
DR EMBL: M26284; AAA37797.1; JOINED.
DR EMBL: M26285; AAA37797.1; JOINED.
DR EMBL: M26286; AAA37797.1; JOINED.
DR EMBL: M26287; AAA37797.1; JOINED.
DR EMBL: M26288; AAA37797.1; JOINED.
DR EMBL: M26289; AAA37797.1; JOINED.
DR EMBL: M26290; AAA37797.1; JOINED.
DR EMBL: M26291; AAA37797.1; JOINED.
DR EMBL: M26292; AAA37797.1; JOINED.
DR EMBL: M26293; AAA37797.1; JOINED.
DR EMBL: M26294; AAA37797.1; JOINED.
DR EMBL: M26295; AAA37797.1; JOINED.
DR EMBL: M26296; AAA37797.1; JOINED.
DR EMBL: M26297; AAA37797.1; JOINED.
DR EMBL: M26298; AAA37797.1; JOINED.
DR EMBL: M26299; AAA37797.1; JOINED.
DR EMBL: M26300; AAA37797.1; JOINED.
DR EMBL: M26301; AAA37797.1; JOINED.
DR EMBL: M26302; AAA37797.1; JOINED.
DR EMBL: M26303; AAA37797.1; JOINED.
DR EMBL: M26304; AAA37797.1; JOINED.
DR EMBL: M26305; AAA37797.1; JOINED.
DR EMBL: M26306; AAA37797.1; JOINED.
DR EMBL: M26307; AAA37797.1; JOINED.
DR EMBL: M26308; AAA37797.1; JOINED.
DR EMBL: M26309; AAA37797.1; JOINED.
DR EMBL: M26310; AAA37797.1; JOINED.
DR EMBL: M26311; AAA37797.1; JOINED.
DR EMBL: M26312; AAA37797.1; JOINED.
DR EMBL: M26313; AAA37797.1; JOINED.
DR EMBL: M26314; AAA37797.1; JOINED.
DR EMBL: M26315; AAA37797.1; JOINED.
DR EMBL: M26316; AAA37797.1; JOINED.
DR EMBL: M26317; AAA37797.1; JOINED.
DR EMBL: M26318; AAA37797.1; JOINED.
DR EMBL: M26319; AAA37797.1; JOINED.
DR EMBL: M26320; AAA37797.1; JOINED.
DR EMBL: M26321; AAA37797.1; JOINED.
DR EMBL: M26322; AAA37797.1; JOINED.
DR EMBL: M26323; AAA37797.1; JOINED.
DR EMBL: M26324; AAA37797.1; JOINED.
DR EMBL: M26325; AAA37797.1; JOINED.
DR EMBL: M26326; AAA37797.1; JOINED.
DR EMBL: M26327; AAA37797.1; JOINED.
DR EMBL: M26328; AAA37797.1; JOINED.
DR EMBL: M26329; AAA37797.1; JOINED.
DR EMBL: M26330; AAA37797.1; JOINED.
DR EMBL: M26331; AAA37797.1; JOINED.
DR EMBL: M26332; AAA37797.1; JOINED.
DR EMBL: M26333; AAA37797.1; JOINED.
DR EMBL: M26334; AAA37797.1; JOINED.
DR EMBL: M26335; AAA37797.1; JOINED.
DR EMBL: M26336; AAA37797.1; JOINED.
DR EMBL: M26337; AAA37797.1; JOINED.
DR EMBL: M26338; AAA37797.1; JOINED.
DR EMBL: M26339; AAA37797.1; JOINED.
DR EMBL: M26340; AAA37797.1; JOINED.
DR EMBL: M26341; AAA37797.1; JOINED.
DR EMBL: M26342; AAA37797.1; JOINED.
DR EMBL: M26343; AAA37797.1; JOINED.
DR EMBL: M26344; AAA37797.1; JOINED.
DR EMBL: M26345; AAA37797.1; JOINED.
DR EMBL: M26346; AAA37797.1; JOINED.
DR EMBL: M26347; AAA37797.1; JOINED.
DR EMBL: M26348; AAA37797.1; JOINED.
DR EMBL: M26349; AAA37797.1; JOINED.
DR EMBL: M26350; AAA37797.1; JOINED.
DR EMBL: M26351; AAA37797.1; JOINED.
DR EMBL: M26352; AAA37797.1; JOINED.
DR EMBL: M26353; AAA37797.1; JOINED.
DR EMBL: M26354; AAA37797.1; JOINED.
DR EMBL: M26355; AAA37797.1; JOINED.
DR EMBL: M26356; AAA37797.1; JOINED.
DR EMBL: M26357; AAA37797.1; JOINED.
DR EMBL: M26358; AAA37797.1; JOINED.
DR EMBL: M26359; AAA37797.1; JOINED.
DR EMBL: M26360; AAA37797.1; JOINED.
DR EMBL: M26361; AAA37797.1; JOINED.
DR EMBL: M26362; AAA37797.1; JOINED.
DR EMBL: M26363; AAA37797.1; JOINED.
DR EMBL: M26364; AAA37797.1; JOINED.
DR EMBL: M26365; AAA37797.1; JOINED.
DR EMBL: M26366; AAA37797.1; JOINED.
DR EMBL: M26367; AAA37797.1; JOINED.
DR EMBL: M26368; AAA37797.1; JOINED.
DR EMBL: M26369; AAA37797.1; JOINED.
DR EMBL: M26370; AAA37797.1; JOINED.
DR EMBL: M26371; AAA37797.1; JOINED.
DR EMBL: M26372; AAA37797.1; JOINED.
DR EMBL: M26373; AAA37797.1; JOINED.
DR EMBL: M26374; AAA37797.1; JOINED.
DR EMBL: M26375; AAA37797.1; JOINED.
DR EMBL: M26376; AAA37797.1; JOINED.
DR EMBL: M26377; AAA37797.1; JOINED.
DR EMBL: M26378; AAA37797.1; JOINED.
DR EMBL: M26379; AAA37797.1; JOINED.
DR EMBL: M26380; AAA37797.1; JOINED.
DR EMBL: M26381; AAA37797.1; JOINED.
DR EMBL: M26382; AAA37797.1; JOINED.
DR EMBL: M26383; AAA37797.1; JOINED.
DR EMBL: M26384; AAA37797.1; JOINED.
DR EMBL: M26385; AAA37797.1; JOINED.
DR EMBL: M26386; AAA37797.1; JOINED.
DR EMBL: M26387; AAA37797.1; JOINED.
DR EMBL: M26388; AAA37797.1; JOINED.
DR EMBL: M26389; AAA37797.1; JOINED.
DR EMBL: M26390; AAA37797.1; JOINED.
DR EMBL: M26391; AAA37797.1; JOINED.
DR EMBL: M26392; AAA37797.1; JOINED.
DR EMBL: M26393; AAA37797.1; JOINED.
DR EMBL: M26394; AAA37797.1; JOINED.
DR EMBL: M26395; AAA37797.1; JOINED.
DR EMBL: M26396; AAA37797.1; JOINED.
DR EMBL: M26397; AAA37797.1; JOINED.
DR EMBL: M26398; AAA37797.1; JOINED.
DR EMBL: M26399; AAA37797.1; JOINED.
DR EMBL: M26400; AAA37797.1; JOINED.
DR EMBL: M26401; AAA37797.1; JOINED.
DR EMBL: M26402; AAA37797.1; JOINED.
DR EMBL: M26403; AAA37797.1; JOINED.
DR EMBL: M26404; AAA37797.1; JOINED.
DR EMBL: M26405; AAA37797.1; JOINED.
DR EMBL: M26406; AAA37797.1; JOINED.
DR EMBL: M26407; AAA37797.1; JOINED.
DR EMBL: M26408; AAA37797.1; JOINED.
DR EMBL: M26409; AAA37797.1; JOINED.
DR EMBL: M26410; AAA37797.1; JOINED.
DR EMBL: M26411; AAA37797.1; JOINED.
DR EMBL: M26412; AAA37797.1; JOINED.
DR EMBL: M26413; AAA37797.1; JOINED.
DR EMBL: M26414; AAA37797.1; JOINED.
DR EMBL: M26415; AAA37797.1; JOINED.
DR EMBL: M26416; AAA37797.1; JOINED.
DR EMBL: M26417; AAA37797.1; JOINED.
DR EMBL: M26418; AAA37797.1; JOINED.
DR EMBL: M26419; AAA37797.1; JOINED.
DR EMBL: M26420; AAA37797.1; JOINED.
DR EMBL: M26421; AAA37797.1; JOINED.
DR EMBL: M26422; AAA37797.1; JOINED.
DR EMBL: M26423; AAA37797.1; JOINED.
DR EMBL: M26424; AAA37797.1; JOINED.
DR EMBL: M26425; AAA37797.1; JOINED.
DR EMBL: M26426; AAA37797.1; JOINED.
DR EMBL: M26427; AAA37797.1; JOINED.
DR EMBL: M26428; AAA37797.1; JOINED.
DR EMBL: M26429; AAA37797.1; JOINED.
DR EMBL: M26430; AAA37797.1; JOINED.
DR EMBL: M26431; AAA37797.1; JOINED.
DR EMBL: M26432; AAA37797.1; JOINED.
DR EMBL: M26433; AAA37797.1; JOINED.
DR EMBL: M26434; AAA37797.1; JOINED.
DR EMBL: M26435; AAA37797.1; JOINED.
DR EMBL: M26436; AAA37797.1; JOINED.
DR EMBL: M26437; AAA37797.1; JOINED.
DR EMBL: M26438; AAA37797.1; JOINED.
DR EMBL: M26439; AAA37797.1; JOINED.
DR EMBL: M26440; AAA37797.1; JOINED.
DR EMBL: M26441; AAA37797.1; JOINED.
DR EMBL: M26442; AAA37797.1; JOINED.
DR EMBL: M26443; AAA37797.1; JOINED.
DR EMBL: M26444; AAA37797.1; JOINED.
DR EMBL: M26445; AAA37797.1; JOINED.
DR EMBL: M26446; AAA37797.1; JOINED.
DR EMBL: M26447; AAA37797.1; JOINED.
DR EMBL: M26448; AAA37797.1; JOINED.
DR EMBL: M26449; AAA37797.1; JOINED.
DR EMBL: M26450; AAA37797.1; JOINED.
DR EMBL: M26451; AAA37797.1; JOINED.
DR EMBL: M26452; AAA37797.1; JOINED.
DR EMBL: M26453; AAA37797.1; JOINED.
DR EMBL: M26454; AAA37797.1; JOINED.
DR EMBL: M26455; AAA37797.1; JOINED.
DR EMBL: M26456; AAA37797.1; JOINED.
DR EMBL: M26457; AAA37797.1; JOINED.
DR EMBL: M26458; AAA37797.1; JOINED.
DR EMBL: M26459; AAA37797.1; JOINED.
DR EMBL: M26460; AAA37797.1; JOINED.
DR EMBL: M26461; AAA37797.1; JOINED.
DR EMBL: M26462; AAA37797.1; JOINED.
DR EMBL: M26463; AAA37797.1; JOINED.
DR EMBL: M26464; AAA37797.1; JOINED.
DR EMBL: M26465; AAA37797.1; JOINED.
DR EMBL: M26466; AAA37797.1; JOINED.
DR EMBL: M26467; AAA37797.1; JOINED.
DR EMBL: M26468; AAA37797.1; JOINED.
DR EMBL: M26469; AAA37797.1; JOINED.
DR EMBL: M26470; AAA37797.1; JOINED.
DR EMBL: M26471; AAA37797.1; JOINED.
DR EMBL: M26472; AAA37797.1; JOINED.
DR EMBL: M26473; AAA37797.1; JOINED.
DR EMBL: M26474; AAA37797.1; JOINED.
DR EMBL: M26475; AAA37797.1; JOINED.
DR EMBL: M26476; AAA37797.1; JOINED.
DR EMBL: M26477; AAA37797.1; JOINED.
DR EMBL: M26478; AAA37797.1; JOINED.
DR EMBL: M26479; AAA37797.1; JOINED.
DR EMBL: M26480; AAA37797.1; JOINED.
DR EMBL: M26481; AAA37797.1; JOINED.
DR EMBL: M26482; AAA37797.1; JOINED.
DR EMBL: M26483; AAA37797.1; JOINED.
DR EMBL: M26484; AAA37797.1; JOINED.
DR EMBL: M26485; AAA37797.1; JOINED.
DR EMBL: M26486; AAA37797.1; JOINED.
DR EMBL: M26487; AAA37797.1; JOINED.
DR EMBL: M26488; AAA37797.1; JOINED.
DR EMBL: M26489; AAA37797.1; JOINED.
DR EMBL: M26490; AAA37797.1; JOINED.
DR EMBL: M26491; AAA37797.1; JOINED.
DR EMBL: M26492; AAA37797.1; JOINED.
DR EMBL: M26493; AAA37797.1; JOINED.
DR EMBL: M26494; AAA37797.1; JOINED.
DR EMBL: M26495; AAA37797.1; JOINED.
DR EMBL: M26496; AAA37797.1; JOINED.
DR EMBL: M26497; AAA37797.1; JOINED.
DR EMBL: M26498; AAA37797.1; JOINED.
DR EMBL: M26499; AAA37797.1; JOINED.
DR EMBL: M26500; AAA37797.1; JOINED.
DR EMBL: M26501; AAA37797.1; JOINED.
DR EMBL: M26502; AAA37797.1; JOINED.
DR EMBL: M26503; AAA37797.1; JOINED.
DR EMBL: M26504; AAA37797.1; JOINED.
DR EMBL: M26505; AAA37797.1; JOINED.
DR EMBL: M26506; AAA37797.1; JOINED.
DR EMBL: M26507; AAA37797.1; JOINED.
DR EMBL: M26508; AAA37797.1; JOINED.
DR EMBL: M26509; AAA37797.1; JOINED.
DR EMBL: M26510; AAA37797.1; JOINED.
DR EMBL: M26511; AAA37797.1; JOINED.
DR EMBL: M26512; AAA37797.1; JOINED.
DR EMBL: M26513; AAA37797.1; JOINED.
DR EMBL: M26514; AAA37797.1; JOINED.
DR EMBL: M26515; AAA37797.1; JOINED.
DR EMBL: M26516; AAA37797.1; JOINED.
DR EMBL: M26517; AAA37797.1; JOINED.
DR EMBL: M26518; AAA37797.1; JOINED.
DR EMBL: M26519; AAA37797.1; JOINED.
DR EMBL: M26520; AAA37797.1; JOINED.
DR EMBL: M26521; AAA37797.1; JOINED.
DR EMBL: M26522; AAA37797.1; JOINED.
DR EMBL: M26523; AAA37797.1; JOINED.
DR EMBL: M26524; AAA37797.1; JOINED.
DR EMBL: M26525; AAA37797.1; JOINED.
DR EMBL: M26526; AAA37797.1; JOINED.
DR EMBL: M26527; AAA37797.1; JOINED.
DR EMBL: M26528; AAA37797.1; JOINED.
DR EMBL: M26529; AAA37797.1; JOINED.
DR EMBL: M26530; AAA37797.1; JOINED.
DR EMBL: M26531; AAA37797.1; JOINED.
DR EMBL: M26532; AAA37797.1; JOINED.
DR EMBL: M26533; AAA37797.1; JOINED.
DR EMBL: M26534; AAA37797.1; JOINED.
DR EMBL: M26535; AAA37797.1; JOINED.
DR EMBL: M26536; AAA37797.1; JOINED.
DR EMBL: M26537; AAA37797.1; JOINED.
DR EMBL: M26538; AAA37797.1; JOINED.
DR EMBL: M26539; AAA37797.1; JOINED.
DR EMBL: M26540; AAA37797.1; JOINED.
DR EMBL: M26541; AAA37797.1; JOINED.
DR EMBL: M26542; AAA37797.1; JOINED.
DR EMBL: M26543; AAA37797.1; JOINED.
DR EMBL: M26544; AAA37797.1; JOINED.
DR EMBL: M26545; AAA37797.1; JOINED.
DR EMBL: M26546; AAA37797.1; JOINED.
DR EMBL: M26547; AAA37797.1; JOINED.
DR EMBL: M26548; AAA37797.1; JOINED.
DR EMBL: M26549; AAA37797.1; JOINED.
DR EMBL: M26550; AAA37797.1; JOINED.
DR EMBL: M26551; AAA37797.1; JOINED.
DR EMBL: M26552; AAA37797.1; JOINED.
DR EMBL: M26553; AAA37797.1; JOINED.
DR EMBL: M26554; AAA37797.1; JOINED.
DR EMBL: M26555; AAA37797.1; JOINED.
DR EMBL: M26556; AAA37797.1; JOINED.
DR EMBL: M26557; AAA37797.1; JOINED.
DR EMBL: M26558; AAA37797.1; JOINED.
DR EMBL: M26559; AAA37797.1; JOINED.
DR EMBL: M26560; AAA37797.1; JOINED.
DR EMBL: M26561; AAA37797.1; JOINED.
DR EMBL: M26562; AAA37797.1; JOINED.
DR EMBL: M26563; AAA37797.1; JOINED.
DR EMBL: M26564; AAA37797.1; JOINED.
DR EMBL: M26565; AAA37797.1; JOINED.
DR EMBL: M26566; AAA37797.1; JOINED.
DR EMBL: M26567; AAA37797.1; JOINED.
DR EMBL: M26568; AAA37797.1; JOINED.
DR EMBL: M26569; AAA37797.1; JOINED.
DR EMBL: M26570; AAA37797.1; JOINED.
DR EMBL: M26571; AAA37797.1; JOINED.
DR EMBL: M26572; AAA37797.1; JOINED.
DR EMBL: M26573; AAA37797.1; JOINED.
DR EMBL: M26574; AAA37797.1; JOINED.
DR EMBL: M26575; AAA37797.1; JOINED.
DR EMBL: M26576; AAA37797.1; JOINED.
DR EMBL: M26577; AAA37797.1; JOINED.
DR EMBL: M26578; AAA37797.1; JOINED.
DR EMBL: M26579; AAA37797.1; JOINED.
DR EMBL: M26580; AAA37797.1; JOINED.
DR EMBL: M26581; AAA37797.1; JOINED.
DR EMBL: M26582; AAA37797.1; JOINED.
DR EMBL: M26583; AAA37797.1; JOINED.
DR EMBL: M26584; AAA37797.1; JOINED.
DR EMBL: M26585; AAA37797.1; JOINED.
DR EMBL: M26586; AAA37797.1; JOINED.
DR EMBL: M26587; AAA37797.1; JOINED.
DR EMBL: M26588; AAA37797.1; JOINED.
DR EMBL: M26589; AAA37797.1; JOINED.
DR EMBL: M26590; AAA37797.1; JOINED.
DR EMBL: M26591; AAA37797.1; JOINED.
DR EMBL: M26592; AAA37797.1; JOINED.
DR EMBL: M26593; AAA37797.1; JOINED.
DR EMBL: M26594; AAA37797.1; JOINED.
DR EMBL: M26595; AAA37797.1; JOINED.
DR EMBL: M26596; AAA37797.1; JOINED.
DR EMBL: M26597; AAA37797.1; JOINED.
DR EMBL: M26598; AAA37797.1; JOINED.
DR EMBL: M26599; AAA37797.1; JOINED.
DR EMBL: M26600; AAA37797.1; JOINED.
DR EMBL: M26601; AAA37797.1; JOINED.
DR EMBL: M26602; AAA37797.1; JOINED.
DR EMBL: M26603; AAA37797.1; JOINED.
DR EMBL: M26604; AAA37797.1; JOINED.
DR EMBL: M26605; AAA37797.1; JOINED.
DR EMBL: M26606; AAA37797.1; JOINED.
DR EMBL: M26607; AAA37797.1; JOINED.
DR EMBL: M26608; AAA37797.1; JOINED.
DR EMBL: M26609; AAA37797.1; JOINED.
DR EMBL: M26610; AAA37797.1; JOINED.
DR EMBL: M26611; AAA3
```

Tue Jul 30 08:40:07 2002

us-09-995-917a-1.rsp

Page 12

```

QY 274 HDLEKEVIVLKGKKVQPIPIFAVYLDLSLHNPENPM-----RWKTTA---FGG 323
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 376 KDTFRDRLTLKRGVEVPMCLSVLAKDPKFSNPDPENPKHFLDDKQGFKSDSAVFVFSIG 435
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 324 VVPCVGGGLGKLTQIAFPLHLYLSYRRK 351
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 436 KVICGEGSLAMDELPLFLITIMQVHFR 463
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: July 29, 2002, 13:58:43
Job time: 70 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 13:57:33 ; Search time 20.97 Seconds
(without alignments)
1750.411 Million cell updates/sec

Title: US-09-995-917A-1
Perfect score: 2027

Sequence: 1 MGMPFGETISFEKPHRSDS.....YVEFKRGMLLEIPTKFLD 382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR71:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2018	99.6	382	2 T48613	hypothetical prote
2	821	40.5	513	2 T46143	steroid 22-alpha-h
3	655	32.3	472	1 S55379	cytochrome P450 CY
4	638	31.5	457	2 D85429	cytochrome P450 11
5	599.5	29.6	512	2 H96759	probable steroid 2
6	568.5	28.0	464	2 T07859	cytochrome P450 ho
7	504	24.9	457	2 T04444	cytochrome P450 -
8	487.5	24.1	485	2 A84859	probable cytochrom
9	484	23.9	482	2 T02739	probable cytochrom
10	460	22.7	455	2 T48973	cytochrome P450-11
11	432.5	21.3	490	2 H86185	hypothetical prote
12	427	21.1	489	2 B84733	probable cytochrom
13	416	20.5	487	1 C71417	cytochrome P450 d1
14	412.5	20.4	460	2 D96813	hypothetical prote
15	390.5	19.3	519	1 T02263	cytochrome P450 DM
16	370	18.3	255	2 T04602	cytochrome P450 ho
17	347.5	17.1	444	1 S75761	cytochrome P450 -
18	330.5	11.4	494	2 A33030	coarnatin 7-monooxy
19	325.5	11.1	494	2 C34271	cytochrome P450 2A
20	223.5	11.0	500	2 A33531	testosterone 15alp
21	223	11.0	494	2 T04737	cytochrome P450 ho
22	219.5	10.8	494	2 B33531	cytochrome P450 2A
23	219	10.8	500	2 T52175	cytochrome P450 mo
24	217.5	10.7	494	2 S16068	testosterone 15alp
25	215.5	10.6	349	2 A86329	FLAPL 4 Protein -
26	215.5	10.6	494	1 O4H0A6	coarnatin 7-hydroxy
27	215.5	10.6	494	2 B47494	cytochrome P450 2A
28	206.5	10.2	494	2 A47494	cytochrome P450 2A
29	204.5	10.1	494	2 A33293	cytochrome P450 2A

ALIGNMENTS

Query Match	Best Local Similarity	Score	2018; DB 2;	Length	382;
Matches	380;	Conservative	1;	Mismatches	1;
Indels	0;	Gaps	0;		
QY 1	MGMPFGETISFEKPHRSDS	IGTFLQQRVSRGKVKFSN	ICGGKAVSCDQELNMFILQ	N	60
DB 1	MGMPFGETISFEKPHRSDS	IGTFLQQRVSRGKVKFSN	ICGGKAVSCDQELNMFILQ	N	60
QY 61	ECKLFTSDYPRKAMHDLGKYS	LLATGTEIHRKLVNII	SFIMLTKSKDPLCAENLSIS	120	
DB 61	ECKLFTSDYPRKAMHDLGKYS	LLATGTEIHRKLVNII	SFIMLTKSKDPLCAENLSIS	120	
QY 121	ILKSQKNCREVEFHKEVKTLS	VNVNQLLSIKRPDPA	RLVYLQDFLSYMKGFISLPI	180	
DB 121	ILKSQKNCREVEFHKEVKTLS	VNVNQLLSIKRPDPA	RLVYLQDFLSYMKGFISLPI	180	
QY 181	PETGTNMIKVSNNRHHQNAL	IIEEDMNNAIREDFLDS	ISNEDEHNAIRAKKGDGELL	240	
DB 181	PETGTNMIKVSNNRHHQNAL	IIEEDMNNAIREDFLDS	ISNEDEHNAIRAKKGDGELL	240	
QY 241	NMEDYQKMEFTQCVISEAL	RCNIVKTYHRRKTHDI	KFEKVEYIPGKMFVFPFTAVH	300	
DB 241	NMEDYQKMEFTQCVISEAL	RCNIVKTYHRRKTHDI	KFEKVEYIPGKMFVFPFTAVH	300	
QY 301	SLHNPPEEFNPWRMTKTA	FAFGGVAVCPGEGELG	KLIQIAFFLHLHLVLSYRMKIKSD	360	
DB 301	SLHNPPEEFNPWRMTKTA	FAFGGVAVCPGEGELG	KLIQIAFFLHLHLVLSYRMKIKSD	360	
QY 361	HPYVEFKRGMLLEIPTKFL	D	382		
DB 361	HPYVEFKRGMLLEIPTKFL	D	382		

RESULT 2
T6143
steroid 22-alpha-hydroxylase (DWF4) - Arabidopsis thaliana
N:Alternate names: protein T3A5.40
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 16-Feb-2001
C:Accession: T6143
R:Blocker, H.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Quetler, F.; Sala
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23024
A:Accession: T6143
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-513 <BL0>
A:Cross-references: EMBL:AL132979
A:Experimental source: cultivar Columbia; BAC clone T3A5
C:Genetics:
A:Map position: 3
A:Intons: 74/2; 182/3; 233/3; 338/3; 369/3; 396/1; 432/3
A:Note: T3A5.40
C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F:308-484/Domain: cytochrome P450 homology <P45>
F:462/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 40.5%; Score 821; DB 2; Length 513;
Best Local Similarity 34.4%; Pred. No. 1.3e-54;
Matches 160; Conservative 87; Mismatches 126; Indels 92; Gaps 4;
QY 2 GWPFIETISFEPKRRSDISGTFLOQVRSYRGVFKSNICGGKAVVSCDDELNMFILQ 61
DB 45 GWPFIETISFEPKRRSDISGTFLOQVRSYRGVFKSNICGGKAVVSCDDELNMFILQ 104
QY 62 GKLETSYPRKAMHDILGKYSLLATGEIHRKLVNIIISFINTLSKDPDLCAENLSIS 121
DB 105 GRLEFCSTYPSIGILGKWSMLVVGDMHRMSISLNFSAHRLRTILKDYERHTLFV 164
QY 122 LKSKNCREVEFHKEVMEFTLSVWVNDLSIKPEDPARLYVLODFLSYMGFSLPIPL 181
DB 165 LKSKNCREVEFHKEVMEFTLSVWVNDLSIKPEDPARLYVLODFLSYMGFSLPIPL 224
QY 182 GTGTNAIKVRSNHNHONAIIEDMNAIREDFLDSISNED----- 224
DB 225 GTATKALQSGATTLKFIERKMEERKLDIKEEDOEVEEVTEDEAEMSKSDHYAKKRTDD 284
QY 225 -----EHAIAI 230
DB 285 DLGLMVLKHSNLSLEQILDLILSLFAGHETSSVALAIFFLQACPAVEELREHLEI 344
QY 231 -RAKKGDEL-LNMEEDYOKMEFTOCVSEALRCGNIVKYTHRKATHDIKFEVYIPKGM 288
DB 345 ARAKKEGESELNWDYKKMDFTQCVINETRLGNVRFHRRKALKDVRKGYDIPSGWK 404
QY 289 VFPIETAVHLPDLSEHNEFEENPRKWTKT-----AFGGVAVCPG 329
DB 405 VLPVTSANHLNDRYDQPNLFNPRWQONNAGASSSGSFTWGNVMPFGGPRICAG 464
QY 330 GELGKQIAFLHLVLSYRKIKSDKEMPIAHPIYVEFRKGMLEI 374
DB 465 SELAKLEMAVFIHHLVLFKNMELAEDDKPFAPFVDPNGDIRV 509

RESULT 3
S55379
cytochrome P450 CYP90 - Arabidopsis thaliana
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: S55379
R:Szekeres, M.; Nemeth, K.; Koncz, Z.; Nagy, F.; Koncz, C.

submitted to the EMBL Data Library, May 1995
A:Reference number: S55379
A:Accession: S55379
A:Molecule type: mRNA
A:Residues: 1-472 <SZR>
A:Cross-references: EMBL:X87367; NID:9853718; PIDN:CA60793.1; PID:9853719
C:Genetics:
A:Gene: CYP90
C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase
F:275-440/Domain: cytochrome P450 homology <P45>
F:418/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 32.3%; Score 655; DB 1; Length 472;
Best Local Similarity 34.5%; Pred. No. 4.7e-42;
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;
QY 1 MGWPIETISFEPKRRSDISGTFLOQVRSYRGVFKSNICGGKAVVSCDDELNMFILQ 60
DB 36 LGIPLIGEFPIGAVKTEPNPEFIDERVARYGVFMHLEFGEPITFSADPETNRFVLQ 95
QY 61 EGKLETSYPRKAMHDILGKYSLLATGEIHRKLVNIIISFINTLSKDPDLCAENLSIS 120
DB 96 EGKLEFCSTYPSIGILGKWSMLVVGDMHRMSISLNFSAHRLRTILKDYERHTLFV 155
QY 121 LKSKNCREVEFHKEVMEFTLSVWVNDLSIKPEDPARLYVLODFLSYMGFSLPIPL 180
DB 156 NIDMSNS--RVLLMEPAKITFELTYKOLMSFDPGEMSE-SLRKEYLVIEGFSPLPL 212
QY 181 PGTYNAIKVRSNHNHONAI-----EDMNAIREDFLDSISNED----- 224
DB 213 FSTYTKAIQAR--RKVALALTVVVKRRREDEEGAEERKMLAALLADDSGFSDEEIVD 270
QY 225 -----EHAIAIRAKKGDELNMEEDYOKME 249
DB 271 FLVALIVAGETTTSTMTLAVKFLTETPLALQKREHRIKRAMSDSLSLMSDYKSM 330
QY 250 FTQCVISEALRCGNIVKYTHRKATHDIKFEVYIPKGMVFPIETAVHLPDLSEHNEFE 309
DB 331 FTQCVNETLRVANIIGVFRRAMTQVEIKGKIPKGMVVFSSFRVHDPNHFIDARTF 390
QY 310 NPKRVRT-----TAFGGVAVCPGGELGKQIAFLHLVLSYRW 350
DB 391 NPKRMSNSVTTGPSNVFTFPGGPRICPGVELARVALSVFLHRLVTGSW 441

RESULT 4
DB5429
cytochrome P450 like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Apr-2001
C:Accession: DB5429
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
Nature 402: 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: AB5001; MIDID:20083488
A:Accession: DB5429
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-457 <STO>
A:Cross-references: GB:NC_001268; NID:g7270586; PIDN:CA80304.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g36380
A:Map position: 4
C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F:396/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 31.5%; Score 638; DB 2; Length 457;
Best Local Similarity 32.2%; Pred. No. 8.8e-41;
Matches 141; Conservative 75; Mismatches 156; Indels 66; Gaps 7;


```

Db 331 RAVILTSRLATLVNGVLTQDMDINGIIPKGRIRIYYRELINDDPLIYPPDYSFNP 390
OY 312 MRWTQTT-----AFGGVAVCPGEGELGKIQIAEFLHLHLVLSYRMKTKSDMPIAHY 363
Db 391 WRWMDKSLHQNSEFLVEGGGTROCPGKELGVAILSTFLHYFYVRYKWEIEGDKLKKFPR 450
OY 364 VEFKRGMLLEI 374
Db 451 VEPANGLIRIV 461

RESULT 7
cytochrome P450 - Arabidopsis thaliana
N:Alternate names: protein T18B16.200; protein T5K18.10
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 28-Jul-2000
C:Accession: T04444; T05806
R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansgorge, W.; Bancroft, I.; Mewes, H.W.
submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15359
A:Accession: T04444
A:Molecule type: DNA
A:Residues: 1-457 <BEV>
A:Cross-references: EMBL:AL021687
A:Experimental source: cultivar Columbia; BAC clone T18B16
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.T.; Voelt, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15453
A:Accession: T05806
A:Molecule type: DNA
A:Residues: 131-457 <BEV>
A:Cross-references: EMBL:AL022580
A:Experimental source: cultivar Columbia; BAC clone T5K18
C:Genetics:
A:Map position: 4
A:Introns: 67/2; 173/3; 302/3; 358/1; 393/3
A:Note: T18B16.200; T5K18.10
C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology
C:Keywords: Chromoprotein; electron transfer; heme; iron; monooxygenase; oxidoreductase
F;272-433/Domain: cytochrome P450 homology <P45>

Query Match 24.9%; Score 504; DB 2; Length 457;
Best Local Similarity 29.3%; Pred. No. 1.3e-30;
Matches 120; Conservative 68; Mismatches 144; Indels 78; Gaps 11;

OY 1 MGMPFGETISFPRKHSDSIGTFLQQRVSRGKVKNSICGKAVVSCDQELNMFILQ 60
Db 41 MGMPYVEETFOY----SDBPNVFQSKOKRGVSFTKTHVLGCPVMISSPEAKFVLVT 96
OY 61 EGGLEFTSDYRKAMHDILGKTSLLANGELHRLKKNVIT--SFT--NLTKSPDFLHCENL 117
Db 97 KSHLFKPTFPASKRMKGKOAIFHQDGYHAKRLKVLRAFMPSIRNMVPI---ESI 152
OY 118 SISILKMKKQREVEPFKEVKMFTLSVMNOLSI-----KPEDPARLYLQDFLSYM 170
Db 153 AODSLRWEGM--INTYQEKTYTFNV---ALLSIFGKDEVLKEDLKRCYIIE----- 203
OY 171 KGFTSLPIPLPGTYTNAIKVRN--RNHQAIIEDMNNAIREDPDLSDIISND----- 224
Db 204 KGYSMPVNLPGTLFHKSMARKELSQLARILISERQNGSSHNDLGSFGDKREELTDE 263
OY 225 -----EHAIRAKKGDGELLNMDY 245
Db 264 QIADNIIGVIFAARDTASVSMILKYLAEINPVLEAVTEQMAIRKDKEGESLTWGD 323
OY 246 QKMFETQCVSEALRCGNIVTVARKAHHDIKFEKVIYIPKGMKVEPIPTAVHLDPSJEN 305
Db 324 KKMPLTRVIOETLRVASITLSFTREAVEDVEYGYLIPKGMKVLPLFRNIHNSADLFSN 383
OY 306 PFEENPMRW-----TKTAFGGGVAVCPGEGELGKIQIAEFLHLHLVLSYR 349

```

```

Db 384 PGKFPDSREVPAPKPTWPMFPFGNGTSCRGNELAKLEMSIMHHLLTKXR 433
RESULT 8
probable cytochrome P450 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 23-Mar-2001
C:Accession: A84859
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MIMD:20083467
A:Accession: A84859
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-485 <STO>
A:Cross-references: GB:AE002093; MID:g4512670; PIDN:AAD21724.1; GSPDB:GN00139
C:Genetics:
A:Gene: Atg42850
A:Map position: 2
C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F;432/Binding site: heme iron (Cys) (axial ligand) #status predicted

```

```

Query Match 24.1%; Score 487.5; DB 2; Length 485;
Best Local Similarity 27.1%; Pred. No. 2.6e-29;
Matches 121; Conservative 79; Mismatches 163; Indels 83; Gaps 10;

OY 1 MGMPFGETISFPRKHSDSI--GTFLQQRVSRGKVKNSICGKAVVSCDQELNMFILQ 59
Db 50 MGMPWGETMDFYKKAOSNVFEDFVNPRLIKGNITKTKMGSPTLVNAGARNLILTS 109
OY 60 NEGKLFSTDYRKAMHDILGKTSLLANGELHRLKKNVIT--ISFINLTKSP-----DF 110
Db 110 NEPSLVVSMWSSSVOLGMNCCIMAKGKHRYLRGVANSLVIGLESILPKICDTVKF 169
OY 111 LHCAENISILSKWKNCRVEERHKEVKMFTLSVMNOLSIKPEDPARLYLQDFLSYM 170
Db 170 HHETE-----WRGKEISLYRSKVLFLVFEPCLYGIVE---TGMLEVEERVL 216
OY 171 KGFTSLPIPLPGTYTNAIKVR-----SNRNHQAIIED 205
Db 217 EGYFALPVEFPCKSFARAKARLEIEFLVGYKREKREMEKGAEPNTTLSRLVEL 276
OY 206 MNNAIREDEFLDSIT-----SNED-----EHAIRAKKGDG 237
Db 277 IKGVITEEVDVNNVLLVFAHDTTSYAMSMTFKMLAQHPTCRDILLQEAHQIKANKGEG 336
OY 238 ELLNMDYQKMEFQCVISRLKCGNIVTVHRAKATHDIKFEKVIYIPKGMKVEPIPTAVH 297
Db 337 EYLTVEDVKKMKYSQWVVRFTMLRSLPIIFSGFRKAVADIDYGGTTPKGMKILMTTYGTH 396
OY 298 LDPSELHNPPEFNPMTQTT-----TAFGGVAVCPGEGELGKIQIAEFLHLHLVLSYRWK 351
Db 397 YNPEIFQDPMSPFTRDKPLQATYIPLFGGGRPLCAGHOLAKSILVFNHFVVTGSDWS 456
OY 352 IKSDMDPIAHYVEF--KRGMLLEIEP 376
Db 457 LVYPDETISMDPLPFPGLGMPKIKSP 482

RESULT 9
T02739
probable cytochrome P450 At2g29090 [imported] - Arabidopsis thaliana
N:Alternate names: cytochrome P450 homolog T914.17
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001
C:Accession: T02739; D84692

```


QY 113 CAENLSTISIKSMKNCREVEFHKEVMTLSVNVNOLL-SIKPEDPARLYLQDFLSYMK 171
DB 147 DVDLARTYMEGARNGYLDVKETSKILLGCLAKKVMGMEPEAKELALCMRY--FQS 204
OY 172 GFISLPRLPGTG-----YTNA-----IKYRSNRNTHQ-- 199
DB 205 GMRFFNLNLTGTGVYKMKVLFVQYTEADISWQARRKMKLKRKTYLTKRASGEELGEFP 264
OY 200 NAIIEDM-----NNAIREDFLDSISNE-----DEEH 227
DB 265 NIFGEMEGEGEMSVENAV-EYIYFFFLVANETPRILLATYKFTSDHKVQGELOREH 323
OY 228 AAI---RAKGDDELLAMEDYQOMETQCVISALRCGNIVTKYHRAKPHDIFKEVYIP 284
DB 324 EEIVRGAAKEGEG--LTWEDYKSMHFQWYINESLRIISTAPYLVRLVLEHDFQVDYTIIP 381
OY 285 KGKVPFPIFTAVHLDPSSLNHPFNPMRW-----TKT-TARGGVRCVPGSELGK 334
DB 382 AGM-FTMGYPIHIFNSEKEDPYAFNPMWKEGDLCAIYKFTIPFGAGRILCVGAFAK 440
OY 335 LQIAFLHLHLVLSYRKIKSDEMPVIAHPYVEFKRMILEI 374
DB 441 MQMAVFIHNL-FRYRMSMKSGTIIIRSFMLMFGGDDVQI 479

RESULT 14
D96813
Hypothetical protein T30F21.17 (Imported) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96813
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federle, N.A.; Kaul, S.; White, O.; Alonso,
Chin, N.C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
Anben, F.F.; Hughes, B.; Hultair, L.
A:Author: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Matzball,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakono, H.
A:Author: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yi, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: D96813
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-460 <SNO>
A:Cross-references: GB:AEO05173; NID:94836883; PIDN:AAD30586.1; GSPDB:GN00141
C:GeneID:8
A:Map position: 17

Query Match 20.4%, Score 412.5; DB 2; Length 460;
Best Local Similarity 27.4%; Pred. No. 1.2e-23;
Matches 121; Conservative 73; Mismatches 163; Indels 85; Gaps 13;
OY 1 MGPPIGTFISFPHNSDSIGTFLOQVRSYRK--VFKSNICGKAVY-----SCDOEL 53
DB 40 MGPPIGTFELDFKPCGVEGIFTEVKKRIRYGRPLFTNIFESKIVSVSTDDVYIHQIFRQ 99
OY 61 ECKLFTSDYPRKAMHDILGKYSLLATGEIHRKLKVAIISFINLTKSKDPFLHCENLSIS 120
DB 100 EMTSEFELGPDIFVAVFGKDNLFLEKVEFIHKYLOKITMOILGSEBCLKTMTGCMNDKATRD 159
OY 121 ILSKMKNCREVEFHKEVMTLSVNVNOLL-SIKPEDPARLY-VLODF-LSYMGFISLP 177
DB 160 HIRTSASGSEFVRKEVENLVAVYMTPLISMLKPEQSKLIDNLAENLDMFKSFLRLS 219
OY 178 IPLPSTGYTNAIKVRSNR-NIHONAIIEDMNAIREDFLDSISNEDE----- 225
DB 220 ---TKKATTKALKSRDEALIQWKVDLMMRKETREKQEDFLVTLLEELKDGSSFDQGSAT 276
OY 226 -----EHAIRAKKGDGEL-LNMEDYO- 246

DB 277 NLIFFALFBRECTSCTLAWEKFTSKDPKVALAEKREKRAIYDNKKDEAGVSEYRN 336
OY 247 KMEFTQCVISEALRCGNIVTKYHRAKPHDIFKE-EVYIRK-----GKVPFIFTAVHLDP 300
DB 337 NMFTMVSNEVLRILANTPLPLFRKAYQDVEIKKTYLEVIMGFGML----- 384
OY 301 SLHENFEENPRHWYTT--AFGGYVRCVPGSELGKLQIAFLHLHLVLSYRKIKSDEMP 358
DB 385 -----WQCKEIMWGSKTMAFGYRLCVGAEFSLQMAFLHLVAVYDFSNQDSEI 438
OY 359 IAHYVEFKRMILEI--EPTK 378
DB 439 IRSPHQYTKDLLINISQSPK 460

RESULT 15
T02263
cytochrome P450 DMARF3 - maize
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Zea mays (maize)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: T02263
R:Winkler, R.G.; Helentjaris, T.
Plant Cell 7, 1307-1317, 1995
A:Title: The maize dwarf3 gene encodes a cytochrome P450-mediated early step in gibbe
A:Reference number: Z14648; MUID:96004534
A:Accession: T02263
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-519 <WIND>
A:Cross-references: EMBL:U32579; NID:9987266; PIDN:AAC49067.1; PID:9987267
A:Experimental source: strain B73
C:Genetics:
A:Gene: dwarf3
A:Function:
A:Description: Involved in an early step in gibberellin biosynthesis
A:Pathway: gibberellin biosynthesis
C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology
C:Keywords: oxidoreductase
F:325-488/domain: cytochrome P450 homology <P45>

Query Match 19.3%, Score 390.5; DB 1; Length 519;
Best Local Similarity 26.4%; Pred. No. 6.6e-22;
Matches 111; Conservative 69; Mismatches 160; Indels 81; Gaps 10;
OY 1 MGPPIGTFISFPHNSDSIGTFLOQVRSYRK--VFKSNICGKAVY-----SCDOEL 53
DB 77 MGWPLVGGMAFLRAFRSGKPDAFIASFVRFRGTGVYRSPFSSPTVLTATAGCKOYL 136
OY 54 NMFILQNEGKLFSDYPRKAMHDILGKYSLLATGEIHRKLKVAIISFINLTKSKDPDLHC 113
DB 137 -----MDDOAFVYGWRKATYALVGPRSFYAMPYDEHRRIRKTLAAINFGDALTGLP 190
OY 114 AENLSTILKSM-KNCEVEFEHKEVMTLSVNVNOLL-SIKPEDPARLYLQ-DFLSYMK 171
DB 191 IDRTVTSRLRAMADHGSEVFLTELRMTFKIYQIFLG--GNDQATTLALERSYTYELNY 248
OY 172 GFISLPRLPGTGCTYNAIKVRSNRNIHONAIIEDMNAIREDE-----DFLDSISNE 224
DB 249 GMRAMAINTLGEFAVYRGLRAR-RLVAVLQGVDERBARAKVSGGCVMMDBLLEAOD 307
OY 225 -----DEHAIRAKK 235
DB 308 ERGRHLDDELDIVLVYLNAGHSSGHITMTATVFQENPDMFAKAKADEALMSIPS 367
OY 236 DGEILNEDYQKMEFTQCVISEALRCGNIVTKYHRAKPHDIFKEVYIPGKAVFIFTA 295
DB 368 SQRLTLRDEPRKMEYLSQVIDETLRLVNISFVSFRQATRDVFNQVGLIPKGMVQVQMYRS 427
OY 296 VHLDPSSLHENPFEENPRHWYTT-----AFGGYVRCVPGSELGKLQIAFLHLHLVLSY 348

Db 428 VMHPQVYPDPPTKFDPSRWEHSPRAGTFLAEGIGARLCPGNDIAKLEISVFLHHFLGY 487
Qy 349 R 349
Db 488 K 488

Search completed: July 29, 2002, 13:58:04
Job time: 31 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 13:57:33 ; Search time 32.67 Seconds
(Without alignments)
1298.751 Million cell updates/sec

Title: US-09-995-917A-1

Perfect score: 2027

Sequence: 1 MCWPFGCTISFFKPHRSDS.....YVEFKRMLETEPTFLIED 382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 aeqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB aeq length: 0

Maximum DB aeq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Database: 1 A_Geneseq_032802.*

1: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being plotted, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	822	40.6	513	21 AAB07921	A cytochrome P450
2	655	32.3	444	21 AAG44572	Arabidopsis thaliana
3	655	32.3	444	21 AAG45023	Arabidopsis thaliana
4	655	32.3	472	18 AAW27153	Arabidopsis thaliana
5	655	32.3	472	21 AAG44571	Arabidopsis thaliana
6	655	32.3	472	21 AAG45022	Arabidopsis thaliana
7	655	32.3	491	21 AAG45021	Arabidopsis thaliana
8	655	32.3	492	21 AAG44570	Arabidopsis thaliana
9	572	28.2	461	21 AAG11836	Arabidopsis thaliana
10	572	28.2	462	21 AAG11835	Arabidopsis thaliana
11	572	28.2	465	21 AAG11834	Arabidopsis thaliana

12	557	27.5	486	22 AAU02839	Taxus cuspidata ox
13	546	26.9	461	21 AAG46490	Arabidopsis thaliana
14	546	26.9	462	21 AAG46489	Arabidopsis thaliana
15	546	26.9	465	21 AAG46491	Arabidopsis thaliana
16	538	26.5	469	21 AAG23014	Arabidopsis thaliana
17	538	26.5	471	21 AAG23013	Arabidopsis thaliana
18	538	26.5	479	21 AAG23012	Arabidopsis thaliana
19	530	25.2	430	21 AAG20784	Arabidopsis thaliana
20	510	25.2	468	21 AAG20783	Arabidopsis thaliana
21	445.5	22.0	500	22 AAU02836	Taxus cuspidata ox
22	432	21.3	493	22 AAU02835	Taxus cuspidata ox
23	427	21.1	481	21 AAG30049	Arabidopsis thaliana
24	427	21.1	489	21 AAG30048	Arabidopsis thaliana
25	425.5	21.0	501	22 AAU02823	Taxus cuspidata ox
26	420	20.7	484	22 AAU02822	Taxus cuspidata ox
27	419	20.7	388	21 AAG20785	Arabidopsis thaliana
28	417	20.6	509	22 AAU02832	Taxus cuspidata ox
29	413	20.4	485	22 AAU02835	Taxus cuspidata ox
30	413	20.4	497	22 AAU02829	Taxus cuspidata ox
31	412.5	20.3	483	22 AAU02827	Taxus cuspidata ox
32	412	20.3	484	22 AAU02821	Taxus cuspidata ox
33	409.5	20.2	507	22 AAU02830	Taxus cuspidata ox
34	408	20.1	512	22 AAU02834	Taxus cuspidata ox
35	402	19.8	496	22 AAU02837	Taxus cuspidata ox
36	396	19.5	433	21 AAG30050	Arabidopsis thaliana
37	387	19.1	500	22 AAU02824	Taxus cuspidata ox
38	381	18.8	503	22 AAU02828	Taxus cuspidata ox
39	342	16.9	492	22 AAU02825	Taxus cuspidata ox
40	339.5	16.7	498	22 AAU02838	Taxus cuspidata ox
41	333.5	16.5	498	22 AAG17905	Arabidopsis thaliana
42	333.5	16.5	242	21 AAG17904	Arabidopsis thaliana
43	333.5	16.5	253	21 AAG17903	Arabidopsis thaliana
44	294.5	14.5	492	19 AAG37733	Cytochrome P450RA
45	294.5	14.5	492	19 AAG44159	Zebrafish retinoid

ALIGNMENTS

RESULT 1
AAB07921 standard. Protein: 513 AA.
XX AAB07921;
XX
XX 14-NOV-2000 (first entry)
XX
XX A cytochrome P450 enzyme designated DWF4.
XX
XX DWF4: cytochrome P450 enzyme; brassinosteroid; 22alpha-hydroxylation;
XX plant phenotype; cell elongation.
XX
XX Arabidopsis sp.
XX
XX WO200047715-A2.
XX
XX 17-AUG-2000.
XX
XX 11-FEB-2000; 2000WO-US03820.
XX
XX 11-FEB-1999; 98US-0119657.
XX 11-FEB-1999; 99US-0119658.
XX
XX (ARIZ-) ARIZONA BOARD OF REGENTS.
XX
XX Azpilro R. Choe S, Feldmann KA;
XX WPI: 2000-549142/50.
XX N-PSDB: AAG59599.
XX
XX New isolated dwf4 polynucleotide useful for altering the phenotype of
XX plants, for diagnostic assays and in the production of antibodies -
XX

PS Claim 50; Fig 11; 113pp; English.

The present sequence represents a DMF4 polypeptide. The polypeptide is a cytochrome P450 enzyme that mediates multiple steps in synthesis of brassinosteroids. Specifically, it mediates multiple 22alpha-hydroxylation steps in brassinosteroid biosynthesis. The DMF4 polynucleotide is used for altering the phenotype of a plant. DMF4 plants display a dramatic reduction in the length of different organs, and this size reduction is attributable to a defect in cell elongation. The DMF4 polynucleotides and polypeptides can be used in diagnostic assays and to generate antibodies, which can be used to produce immunogenic compositions.

Sequence 513 AA;

Query Match	40.6%	Score 822	DB 21	Length 513
Best Local Similarity	34.4%	Pred. No. 2.2e-70		
Matches 160; Conservative	87	Mismatches 126	Indels 92	Gaps 4

[illegible]

RESULT 2

ID AAG44572 standard; Protein; 444 AA.

AC AAG44572;

DT 18-OCT-2000 (first entry)

AA	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 55848.
AA	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 55848.

AA Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter
KW termination sequence.

05 *Arabidopsis thaliana*.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.
VY

[illegible]

PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142917.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145219.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155569.

PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 22-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 32.3%; Score 655; DB 21; Length 444;
Best Local Similarity 34.5%; Pred. No. 2, 2e-54;
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

QY 1 MGWPFGEITSEFFKRRSISIGFLQQRVSRKGVKFSNICGKAVVSCDQELMNEFIION 60
DB 8 IGIPIIGETfqlIgaYktenpepfideravarysvfemthlfgeptlfssadpetnrfvIqn 67
QY 61 EGGLETSDPKAMHDITLGKVSLLATGETHRRKLNKNTISFINLTSKRPPLCAENLSTS 120
DB 68 egklfecsypastcniIghhsllImkgsIhkmhslmfsfianssIkhdlmIdrlvrf 127
QY 121 ILSMKNCREVEFEHKEVKMETLSVMVNOILSTIKPEDPARLVYLODFLSYMKGFISLPIPL 180
DB 128 nIdswas--tvllmeaakkIfelcvkqlmsfdpgwase-slrkeyllvIegffsIplrl 184
QY 181 PGTGYNAIKVSNRNRIHONAIT-----EDMNNAIRREDLDSITSNED----- 224
DB 185 fctLykaiIgar-irkvaalcvvmvkmkreeseegaeirkkmlaallaaIdgffsdeelyd 242
QY 225 -----EEHNAIRAKKGDDGLMNEPYOKME 249
DB 243 flvallvagyettstclmIavkflecprIalagIkeehketrIamsdsysIewsdYkmp 302
QY 250 FTQCIVISEALRCGNITVKVTHRRATHDIKREKXVIRPKGMKVPRTPTAVHLDPSLHBNPEEF 309
DB 303 ftcgvneclIvanIlgvfframcdveIkykIrkYkwvfasfravhIdpnhfkIdartf 362
QY 310 NPMRMTKT-----TAFGGVRCVRCGSGELKQLQIAFHHHLVLSRW 350
DB 363 nPwtgmsnsvltcpnsnvftPrg99pIrlcpgyelaIvalsvfIhrlvtgfsaw 413

RESULT 3
AAG45023
ID AAG45023 standard; Protein; 444 AA.

```
XX AC AAG45023;
XX 18-OCT-2000 (first entry)
DE DE Arabidopsis thaliana protein fragment SEQ ID NO: 56470.
XX XX
XX Protein Identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129645.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 07-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
```

```
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 32.3%; Score 655; DB 21; Length 444;
Best Local Similarity 34.5%; Pred. No. 2.2e-54;
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

```
OY 1 MGWPIGETISFFKPHRSDSIGTFLQQRVSRYGKVFNSNICGKAVVSCDOELNMFILON 60
DB 8 lglpllgctfqlgkyteknpepfldervarygvmfchlfgcpilfaadpctrnfvlqn 67
OY 61 ECKLTSOYPRAMHDIILKYSLLATGELIRKLRKLVIISEFNLTNKSXPDPLHCAENNISIS 120
DB 68 egklteceypaicalnllqkshlllmkgsllhkrmslmsfnsllkdhlmldldrlvrf 127
OY 121 ILKSWKNCREVEFHKREVKMFTLSVMNQLSIKPEDPARLVVLODFLSYMGFTSLPIPL 180
DB 128 nldeaws--rrllmeeakktifeltvklmsfdpgewse-sirkeylliviegftslppl 184
OY 181 PGTGTTNAIKVRSNRNIIQMAII-----EDMNNAIREDDFLSDISINED----- 224
DB 185 fstlyrkaigqr--rkvealtvwmkrreeeegeaerkkdmlaallaaaddgfsdeetlvd 242
```

```
OY 225 -----EENAIIRAKKGDELLNWDYQKE 249
DB 243 flvalvaygetstlmclavfltecpialaqllkeehckiramdsyslwsdykemp 302
OY 250 FTQCIVISEALRCGNIVKTVRKATHDIRKREYVIRKGNKVFPIFMAVHLDPSLHNPREF 309
DB 303 ftgcgvnetlrvanllggyfirramcdvelkykllpkywkvfsstfavnldphnfkdartf 362
OY 310 NPMRWTKT-----TAFGGVAVCPGCGELGRIQIAFELHNLVLSYRW 350
DB 363 nprwqsnsvttgpanvfcpfgg9prrlcpgyelarvalsvfhrivtgfsw 413

RESULT 4
AAW27153
ID AAW27153 standard; Protein; 472 AA.
XX
AC AAW27153;
XX
DT 14-APR-1998 (first entry)
XX
DE Arabidopsis thaliana cytochrome P450-type hydroxylase.
XX
KM Cytochrome P450-type hydroxylase; identification; brassinosteroid;
KM brassinosteroid inhibitor; modified plant; recombinant production;
KM teasterone.
XX
OS Arabidopsis thaliana.
XX
PN M09735986-A1.
XX
PD 02-OCT-1997.
XX
PE 27-MAR-1997; 97WO-EP01586.
XX
PR 27-MAR-1996; 96US-0622166.
XX
PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
PI Altmann T, Koncz C, Mathur J, Szekeres WA;
XX
XX WPI: 1997-489649/45.
XX DR N-PSDB; AAT85306, AAT85307.
XX
PT New isolated plant cytochrome P450-type hydroxylase gene - used to
PT identify substances acting as brassino-steroid(s) or brassinosteroid
PT inhibitors for the production of modified plants
XX
PS Claim 1; Pages 44-46; 77pp; English.
XX
CC The present sequence is Arabidopsis thaliana cytochrome
CC P450-type hydroxylase. The hydroxylase can be used to identify
CC brassinosteroids or brassinosteroid inhibitors, useful to produce
CC plants with modified physiological and/or phenotypic
CC characteristics. The modified plants may show, e.g. stimulation of
CC growth, increased seed elongation, increased wood production,
CC accelerated seed germination at low temperatures, an increase in
CC dry weight, repressed anthocyanin production during growth in light
CC and/or inhibited de-etiolation which is induced, e.g. by cytokinin,
CC in the dark or an increase in stress tolerance. The hydroxylase or
CC its coding sequence can also be used for the recombinant production
CC of compounds, e.g. teasterone.
XX
SQ Sequence 472 AA;
```

Query Match 32.3%; Score 655; DB 18; Length 472;
Best Local Similarity 34.5%; Pred. No. 2.4e-54;
Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

```
OY 1 MGWPIGETISFFKPHRSDSIGTFLQQRVSRYGKVFNSNICGKAVVSCDOELNMFILON 60
DB 185 fstlyrkaigqr--rkvealtvwmkrreeeegeaerkkdmlaallaaaddgfsdeetlvd 242
```


PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	03-AUG-1999;	99US-0146389.
PR	04-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	05-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	06-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	11-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	23-AUG-1999;	99US-0149923.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159283.
PR	13-OCT-1999;	99US-0159284.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159684.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160960.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	25-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match	32.3%.	Score 655; DB 21; Length 472;
Best Local Similarity	34.5%.	Pred. No. 2,4e-54;
Matches 142; Conservative 70; Mismatches 133; Indels 66; Caps		
OY	1	MGMPIGETISFEFKRRSDSISTGTFLOQRVSRGVFKNSNICGGKAVNYSODELNMFIQON 60
Dd	36	IglrligetfqlgaYktempepfidervarysvfmhlgeptlsadpetnrtivqn 95
OY	61	EGLKFTSDYPKAMHDILGKYSLSLATGEIHRKLKVIIISFINTLRSPDFLCAENLSIS 120
Dd	96	egklfecypasicnllgkbhlllmkgslhkrmslmsfnasslkdhlnldidrvrf 155
OY	121	ILSKWKNREVEFHKEVMFTLSVYNOLSIKEPDPARLVLODPLSYMKGFISLPPL 180
Dd	156	nldsws--rylmeeaekklftelvtqklsfdcpewse-slrkeylvilegfsslppl 212
OY	181	PGTGYNAIKYRNSNRNIHQNAII-----EDMNNNAIREEDFLDSIISNEE----- 224
Dd	213	fstcykralgr--rkvaeatltvymkrreeeeagaekkkmlaallaaodgfsdeatvd 270
OY	225	-----EEHAIRAKKGDELLNWEDYOME 249
Dd	271	flvallyagyetstlmclavkfllteprlaqlkeehkrlranksdslewedykmp 330
OY	250	FTCGVISEALRCGNIVTVHRRKHADHKFEKXYAPKGMKVPRTTAVHLDPSSLHENPEF 309
Dd	331	ft

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138099.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151428.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157173.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158322.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.

OY 310 NPMKMTT-----TAFGGVRYVPCGELGKLIQIAFFLHHLVLSYRM 350
| | | | | : | | | | : | | | | : |
Db 410 nprwtgnsvtltgpnsvfltpfgsgprltcpgyrelarvalsvflhrjvtgfsfw 460

RESULT 8
ACG4570
ID AAG44570 standard; Protein; 492 AA.
XX
XX AAG44570;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 55846.
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EP103405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130444.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145513.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

```
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 32.3%; Score 655; DB 21; Length 492;

Best Local Similarity 34.5%; Pred. No. 2,6e-54;

Matches 142; Conservative 70; Mismatches 133; Indels 66; Gaps 6;

```
QY . 1 MGWPFIEGTTSEFKPHRSDSIGTFLOQRVSRKGVKFSNIGCGKAVYSCQOELMFTLN 60
Db 56 1gdlpigtetqllgayktenpepidervarysvfmthlfgpeplfsadpetnrtvlgp 115
QY 61 EGRKFTDPRAMHDIIGKXSLILATGEIIRKLNVIISFNLTKRSKPDFLHCAENLSIS 120
Db 116 egklfecsyasclnllgkshllmksglnkrmshlmstamfansiikchlmldidrlvrf 175
```

```
QY 121 TLKSMKNCREVEERHKEYKMTLSVMYNOLLSIKPEDPARLVYLQDLSYMKGFISLPIPL 180
Db 176 nldswss--rvllmeeakkltfelvqqlmsfjdpgeuse-slkexyllyiegffsfpjpl 232
QY 181 PGTYTNAIKVRSNRNIRHQAII-----EDMNNAIREEFDLDSIISNED----- 224
Db 233 fsttyrkaigqr--rkvaeltvymkrrreegeegarkkdmiaallaaddgfsdeeiivd 290
QY 225 -----EENHAIRAKKGDGELLNEDYOKME 249
Db 291 flvalliwaqyettslmtlavkflteplalaqknehekrlmksdysyslwsdykmp 350
QY 250 FTQCVISEALRCGNIVKTVHRKATHDIKEKEYVYIPKGWVPIFTAVHLDPSLHNPFEF 309
Db 351 ftgcgvnetlrvaniigvfrimtdveikikpyskvsftravhldpnhfkdarft 410
QY 310 NPMKMTKT-----TAFGGGVRCPPGEGELCKIQIAFFLHLVLSYRW 350
Db 411 npwrwgsnsvltgspnsvfthfgggrplcpgyelarvalsvflhrlytgfsw 461
```

RESULT 9

AAgl1836 standard; Protein; 461 AA.

AAgl1836;

17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 10708.

Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

Arabidopsis thaliana.

Ep1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

```
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
```

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR	19-JUL-1999;	99US-0144333
PR	19-JUL-1999;	99US-0144334
PR	19-JUL-1999;	99US-0144335
PR	20-JUL-1999;	99US-0144352
PR	20-JUL-1999;	99US-0144632
PR	20-JUL-1999;	99US-0144884
PR	21-JUL-1999;	99US-0144814
PR	21-JUL-1999;	99US-0145086
PR	21-JUL-1999;	99US-0145087
PR	22-JUL-1999;	99US-0145088
PR	22-JUL-1999;	99US-0145089
PR	22-JUL-1999;	99US-0145192
PR	23-JUL-1999;	99US-0145145
PR	23-JUL-1999;	99US-0145218
PR	23-JUL-1999;	99US-0145224
PR	26-JUL-1999;	99US-0145276
PR	27-JUL-1999;	99US-0145913
PR	27-JUL-1999;	99US-0145918
PR	27-JUL-1999;	99US-0145919
PR	28-JUL-1999;	99US-0145951
PR	02-AUG-1999;	99US-0145386
PR	02-AUG-1999;	99US-0146388
PR	02-AUG-1999;	99US-0146389
PR	03-AUG-1999;	99US-0147038
PR	04-AUG-1999;	99US-0147204
PR	05-AUG-1999;	99US-0147302
PR	05-AUG-1999;	99US-0147192
PR	06-AUG-1999;	99US-0147260
PR	06-AUG-1999;	99US-0147303
PR	09-AUG-1999;	99US-0147416
PR	09-AUG-1999;	99US-0147493
PR	09-AUG-1999;	99US-0147935
PR	11-AUG-1999;	99US-0148171
PR	11-AUG-1999;	99US-0148319
PR	12-AUG-1999;	99US-0148341
PR	13-AUG-1999;	99US-0148565
PR	13-AUG-1999;	99US-0148684
PR	16-AUG-1999;	99US-0149368
PR	17-AUG-1999;	99US-0149175
PR	18-AUG-1999;	99US-0149426
PR	20-AUG-1999;	99US-0149722
PR	20-AUG-1999;	99US-0149723
PR	23-AUG-1999;	99US-0149929
PR	23-AUG-1999;	99US-0149902
PR	23-AUG-1999;	99US-0149930
PR	25-AUG-1999;	99US-0150566
PR	26-AUG-1999;	99US-0150884
PR	27-AUG-1999;	99US-0151065
PR	27-AUG-1999;	99US-0151066
PR	30-AUG-1999;	99US-0151080
PR	31-AUG-1999;	99US-0151303
PR	01-SEP-1999;	99US-0151948
PR	07-SEP-1999;	99US-0151930
PR	10-SEP-1999;	99US-0152363
PR	13-SEP-1999;	99US-0153070
PR	15-SEP-1999;	99US-0153758
PR	16-SEP-1999;	99US-0154018
PR	20-SEP-1999;	99US-0154039
PR	22-SEP-1999;	99US-0154779
PR	23-SEP-1999;	99US-0155139
PR	24-SEP-1999;	99US-0155486
PR	28-SEP-1999;	99US-0155686
PR	29-SEP-1999;	99US-0156458
PR	04-OCT-1999;	99US-0157117
PR	05-OCT-1999;	99US-0157753
PR	06-OCT-1999;	99US-0157865
PR	07-OCT-1999;	99US-0158029
PR	08-OCT-1999;	99US-0158232
PR	12-OCT-1999;	99US-0158369
PR	13-OCT-1999;	99US-0159293
PR	13-OCT-1999;	99US-0159294

PR	13-OCT-1999;	99US-0159329;
PR	14-OCT-1999;	99US-0159329
PR	14-OCT-1999;	99US-0159330
PR	14-OCT-1999;	99US-0159331
PR	14-OCT-1999;	99US-0159337
PR	14-OCT-1999;	99US-0159368
PR	18-OCT-1999;	99US-0159584
PR	21-OCT-1999;	99US-0160741
PR	21-OCT-1999;	99US-0160767
PR	21-OCT-1999;	99US-0160768
PR	21-OCT-1999;	99US-0160770
PR	21-OCT-1999;	99US-0160814
PR	21-OCT-1999;	99US-0160815
PR	21-OCT-1999;	99US-0160815
PR	22-OCT-1999;	99US-0160980
PR	22-OCT-1999;	99US-0160981
PR	22-OCT-1999;	99US-0160981
PR	25-OCT-1999;	99US-0161404
PR	25-OCT-1999;	99US-0161405
PR	25-OCT-1999;	99US-0161406
PR	26-OCT-1999;	99US-0161359
PR	26-OCT-1999;	99US-0161360
PR	26-OCT-1999;	99US-0161361
PR	26-OCT-1999;	99US-0161920
PR	28-OCT-1999;	99US-0161920
PR	28-OCT-1999;	99US-0161921
PR	29-OCT-1999;	99US-0162142
PR	29-OCT-1999;	99US-0162143

Query Match	28.2%;	Score 572;	DB 21;	Length 462;
Best Local Similarity	30.7%;	Pred. No. 2.4e-46;		
Matches 136;	Conservative 71;	Mismatches 152;	Indels 84;	Gaps 9;

```

OY      1  MGNWPIGTSIFPMRHNBSJCTGLQDRVSRKGVKVPVNSJGCGAAVYSCDOLNMFLO 60
Db      36  mwmrlfsgctceflkygrd-----fmknqgrlrygsfkskhlqgpcrlsvmaaealnrylhm 90
OY      61  ECKLETSDYPRAMHDILKYSYLLATATEIHRKLNKVIISFINTLTKSPDFLHCACENLSIS 120
Db      91  eskvgvaypgsdmdldgtcnlaevhgrshlmzgsllslsptcmkchllprkldtdmtn 150
OY      121  ILKSKNCRREVEFEKVKMFMFLSWVNOQLSI-----kpe-DPARLVYLODLSTYMKGFT 174
Db      151  ylcgwddetevdqektkhm-----afissllqaecltkpweevy-----tefkilvgtl 203
OY      175  SLPIRLPCTGTNAIKRVSNNNHQAIIEDMNAIKKE-----DFLSIINED----- 224
Db      204  svyldlpgtlyrsgvgrmnlldrltlelme-----rkesgetctmdlylmkkehndryl 258
OY      225  -----EENAAIRAKKGDELLN 241
Db      259  ltkelldqvcllylsgyevstsvmaalkyldhnrkaleelrtrethalektrpreptl 318
OY      242  WEDQOKMEFPQCVSEALRSGNIYKTVNRKATNDIKKEVYIRPKGMKVRFIPAVNHLDS 301
Db      319  lddksmkltcaavlfesrlactlvpnrlfcthdlelngllrkyvrluyvtrelnrldys 378
OY      302  LHENPEEFNPRKWTKT-----AFGCGVAVRSCPGELSLQIAEFLNLYLVSYWKIK 353
Db      379  lyedgmflfnprwmekeleeskyfllfsgvrlcprkelyaevsflnyfuytkuyreen 438
OY      354  SDEMIRINPRYERKRGKULETER 376
Db      439  gedklnwfrgvssapkyhllcsp 461

```

```

RESULT 11
AA011834
ID   AA011834 standard; Protein; 465 AA.
AC
XX
XX   AA011834:
XX
XX   17-OCT-2000 (first entry)
XX
XX   Arabidopsis thaliana protein fragment SEQ ID NO: 10706.

```

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hydridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141844.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.


```

PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161361.
PR 26-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

```

Query Match 28.7%; Score 572; DB 21; Length 465;

Best Local Similarity 30.7%; Pred. No. 2,4e-46; Matches 136; Conservative 71; Mismatches 152; Indels 84; Gaps 9;

```

QY 1 MGMPFGETISFFKPHRSDSICGTFLOQRVSRGKVFKSNICGKAVVSCDQELMNFILON 60
DB 39 mgwplfgtcteflkgpdp-----lmknglrlygseffkshllgcpctisvmdelnrrylm 93
QY 61 ECKLFTSDYPRKAMHDILGKYSLLATGTHRKLNKVNIIISFTNLTKSKPDLHCAENLSIS 120
DB 94 eskglvgyppsmldlglctclaaehvshrlmrgslslslspcmkdhllpkldfmrn 153
QY 121 ILKSKMKNCREVEFFKHEVFMFLSVWVNDLSI-----KPE-DPARLYVLODFLSYMGCFI 174
DB 154 ylgcwddleevdldqekthm---afllslqlaelklkpeveyr---teffklvvgtl 206
QY 175 SLPLPCTGYTNAIKVSNRNHONAIIEDMNNNAIREE-----DFLDSIISNED----- 224
DB 207 svpdlipstnyrgvqarndldrlteltmge-----tkesgetftdmlylmkkehndryl 261
QY 225 -----BEHNAIRAKKGDELIN 241
DB 262 ltdkeitdqvvtlllysgyetvstsmmalkylhphkpaeeelrrehlahrerkpdpelc 321
QY 242 WEDYQKMEFTQCVISEALRCGNIVKTVHRAKATHDKIKREYIIPGKVVPIFTAVHLDPS 301

```

```

DB 322 lddiksmkfrvileterlativngvrlrcthdlelngyllpkwrrlyvvtreinydts 381
QY 302 LHENPEFENPMRKTKT-----AFGCVAVCPGCELGKQLIAFFLHHVLVSRRKIK 353
DB 382 lyedpmlfnprwmeksleskeyfllfggyvrlcpqkglgslsevsflhyfkytween 441
QY 354 SDEMIAHPYVEFRKGMLEIEP 376
DB 442 gedklmvtprvsapkgyhlkcs 464

```

RESULT 12

```

ID AAU02839 standard; Protein: 486 AA.
XX AAU02839:
XX 07-SEP-2001 (first entry)
XX Taxus cuspidata oxygenase enzyme #19.
XX Taxus cuspidata oxygenase enzyme #19.
XX Oxygenase: Japanese yew; Taxol; taxoid; Taxol biosynthetic pathway;
XX transgenic organism.
XX Taxus cuspidata.
XX WO200134780-A2.
XX 17-MAY-2001.
XX 13-NOV-2000; 2000WO-US31254.
XX 12-NOV-1999; 9905-0165250.
XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX Croteau RB, Schoendorf A, Jennewein S;
XX WPI: 2001-355489/37.
XX N-PSDB: AAS05196.
XX Novel nucleic acid and amino acid sequences, isolated from the Taxus
XX pt genus, useful for the synthetic production of Taxol and related
XX taxoids, intermediates within the Taxol biosynthetic pathway, and
XX other taxoid derivatives.
PS Claim 1; Page 139-141; 141pp; English.
XX The sequence represents a Taxus cuspidata oxygenase, encoded by a DNA of
XX the invention. The sequences isolated from the Taxus genus, and the
XX respective oxygenases are useful for the synthetic production of Taxol
XX and related taxoids, as well as intermediates within the Taxol
XX biosynthetic pathway, and other taxoid derivatives. The sequences also
XX can be used to make transgenic organisms that either produce the
XX oxygenases for subsequent in vitro use, or produce the oxygenases in vivo
XX so as to alter the level of Taxol and taxoid production within the
XX transgenic organism. The oxygenase nucleic acids and amino acids are
XX useful for isolating the polynucleotide and polypeptide sequences
XX corresponding to full-length oxygenases.
XX Sequence 486 AA;

```

Query Match 27.5%; Score 557; DB 22; Length 486;

Best Local Similarity 30.9%; Pred. No. 7.1e-45; Matches 126; Conservative 64; Mismatches 158; Indels 60; Gaps 8;

```

QY 2 GMPFGETISFFKPHRSDSICGTFLOQRVSRGKVFKSNICG-KAVVSCDQELMNFILON 60
DB 52 glpfigetlftldakpgrtrkffidelnrltyspdlfcsllgtrtravsvpeltkxvlpn 111
QY 61 ECKLFTSDYPRKAMHDILGKYSLLATGTHRKLNKVNIIISFTNLTKSKPDLHCAENLSIS 120

```

Db 112 egrifesaalpfnljgkyisavgeigrklnatavnlkhetssdmedidqda 171
QY 121 ILKSWKRCREVEFHKEVMEFTLSVWVNOILSIKP-EDPARLYVLQDFLSYMKGFSLPIP 179
Db 172 gmrkweegddipbqkncqivlnlmakrlldlpseemghly--kafddfgavlsfjln 229
QY 180 LPTGYTNAIKVR-----SNRNHONAIED-MNNAIREDDLDSTISME---- 223
Db 230 lpgtlyargitarglllkrllkcklrethpevlnldltkivregtsdeladtlilff 289
QY 224 -----DEPNAIRAKKGDG-ELLMWEDYOKMEFTQC 253
Db 290 vfaygetaamantfaykylaenpraleelraehdallkakygnekltwmdyqsmkfvhc 349
QY 254 VISEALRCGNITKVTYHRAKATHDIKREXYIPKGMKFPDIFTAVHLDPSLHENPEFENPMR 313
Db 350 vlnetlrltgatvvlfrreakqdkvkdfvlpkgytvsflsathdgykhyeadkflpwr 409
QY 314 WTKT-----TAFGGVAVCPGGELGKIQIAFELHHLVLSYRWK 351
Db 410 wneqgetleecpymfgrggrlcpjlnarfeialflnhtvktlwe 457
RESULT 13
AAG46490
ID AAG46490 standard; Protein: 461 AA.
XX AAG46490;
AC
XX
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SRQ ID NO: 58494.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132887.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145921.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150586.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161932.
PR 29-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 26.9%; Score 546; DB 21; Length 461;
Best Local Similarity 29.7%; Pred. No. 7.6e-44;
Matches 130; Conservative 73; Mismatches 101; Indels 74; Gaps 9;

QY 1 MGPPIGCTISFEFKPHRSDSICFLQQRVSRGKVFKNICGKAVVSCDGLMNFILQ 60
DB 35 mwpflfgetetclkgpn-----fmnqrlrygsffkehlhgpcpllamdeevnrylkn 89
QY 61 ECKLFTSDVPKAMHDLICKYSLLATGELHKKLVKAVIISFINLTKSRKDFLHCENLSIS 120
DB 90 eakglvpygpgsmldllqtcmaavhgsbhlmrqslslsstmrmchllpkvdfmrs 149
QY 121 ILKSNKREVEFEHKEVKMFTLSVMVNOQLS--IKPEDPARLYLQ--DFLSYMKCFIS 175
DB 150 yldqnelevidqdktkhmalflslltqlagnlrp-----fveefktaffklvgtls 203
QY 176 LPILPGTCGYTNAIKVRSNRNIHQNAIIEDMNA-----IREEDF 215
DB 204 vpldlpgtnyrcgldqarnldfllrelmqerrdsgetfcldmlylmkkggnrypltdael 263
QY 216 LDSIIS-----NEDEHNAIRAKKDGCELNMWEDYOK 247
DB 264 rdqavvlllysgyevstcmmalkylhdpkalgelraehlafrckrqdelpgldevks 323
QY 248 MEFTQCVISEALRCGNIVKTVHRAKATHDIKFEKVVYIPKGMKVPFETVAHLDPSLHNEF 307
DB 324 mkftravlyetarlavlgvrlktrdrelngylipkgyrlyvvtrelngdanlyedpl 383
QY 308 EENPMRWTKTT-----AFGCGVAVCPGGLGKQIAFPFLHNLVLSRW-KISDENP 358
DB 384 lfnpwrmkkslesqnsctvfqggrlcpqkelgylvelssflhyfvttryweelsgdel 442
QY 359 IAHPYVEFKRCMLLEIEP 376
DB 443 mvfprvfapkgfhlrlsp 460

RESULT 14
AAG46489
ID AAG46489 standard; Protein; 462 AA.
XX
XX AAG46489;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58493.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
PN EPI033405-A2.
PD
XX
XX 06-SEP-2000.
PF
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128274.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135184.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142554.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148655.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151348.
PR 07-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160988.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 26 98; Score 546; DB 21; Length 462;
Best Local Similarity 29.78; Pred. No. 7,6e-44;

Matches 130; Conservative 73; Mismatches 161; Indels 74; Gaps 9;

QY 1 MGWPIGCTISFFKPHRDSIGTFLOQRVSRGKVFKSNICGKAVVSCDOELMNFILQN 60
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
Db 36 mgwpi gct i s f f k p h r d s i g t f l o q r v s r g k v f k s n i c g k a v v s c d o e l m n f i l q n 60
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
QY 61 ECKLFTSYPRKAMHDLGKYSLLATGELHKKLVNITSLTKSKEDFLHCAENLSTIS 120
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
Db 91 esk l f t s y p r k a m h d l g k y s l l a t g e l h k k l v n i t s l t k s k e d f l h c a e n l s t i s 120
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
QY 121 ILKSWKNREVEYFHKVEYMFSLVWVNOJLS--IKREPALVYUQ--DPLSTMKGFIS 175
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
Db 151 y l d q n e l e v i d l g d k t c m a a v h g s h l m r g s l l s l s t e m r c h l l p k v d h f m r s 150
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
QY 176 LPIPLPGTGYTNAIKVRSNRNIHQNAILEDNMNA-----IREEDF 215
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
Db 205 v p i d l p g n y r c g i g a r n n i d r l l r e l m g e r d g e t c d m l g y l m k e g n r y p i d e e l 264
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
QY 216 LDSIIS-----NEDEHAAIRAKKGDELNMEDYOK 247
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
Db 265 r d q v n t l i y s g e l v e t l e m m a l y l n d h p k a l g e l r a e h l a f e r e k r q d e p l g l e d v k s 324
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
QY 248 MEFTQCVISELALRCGNIVTKVHRKATDIKREKVVIRPGWKVVFRTFAVNHDPFLHENPF 307
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
Db 325 m k f c a v i y e s r i a c t l v n g l r k t c r d l e n g l l p r g w l i y u y r e i n d a n l y e p l 384
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
QY 308 EFNPMRWTKYT-----AFGGGVRCVPCGSELGLOIAFFLNHLVLSRYR-KIKSDMP 358
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
Db 385 l f n p w r m k k s l e a g n s c f v f g g t r l o p g k e l j v e l s e l t h y f v r l r y w e e l g g e l 443
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
QY 359 IAHRYVEKRGMLLEIEP 376
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |
Db 444 m v f d r v l a p k g f h l r l a p 461
| | | | | : : : | | | | : : | | | : | | : | | : | | : | | : |

RESULT 15
AAG46491
ID AAG46491 standard; Protein; 465 AA.

XX AAG46491;
AC
XX 18-OCT-2000 (first entry)

XX XX Arabidopsis thaliana protein fragment SEQ ID NO: 58495.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
RV termination sequence.
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
PF
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.

PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	26-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	02-JUL-1999;	99US-0142154.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	04-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.

PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	26-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160776.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160960.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match 26.9%; Score 546; DB 21; Length 465;
Best Local Similarity 29.7%; Pred. No. 7.7e-44;
Matches 130; Conservative 73; Mismatches 161; Indels 74; Gaps 9;

QY	1	MGMPTGETISFHKHRDSTIGTFLQOVRSGYKVKSGNICGKAVSCDDELNMFIION	60
DB	39	mgwplfgetelkkgpn-----fmnqrlllygsfkskhllygcpflismdevmrlylkn	93
QY	61	EGKLFSTDPKAMHDILKYSYILATGEIHRKLKVVIIISFTNLTKSKDPFLCAENLIS	120
DB	94	eskglvpyppgsmidlilgtcmaavhgshtlmrgslslslsttmrhdhlpkvdhfmrs	153
QY	121	ILKSKNCREVEFHKVKWFTLSVWVNOILS--IKPEDPARLYVUO---DLSLYKGFIS	175
DB	154	yldqwelevldigdktkmalfsslqtqagnlkrp-----fyeeftatfkilvgtls	207
QY	176	LPRIPLPGTYTNAIKVRSRNRHQAIIEDMNA-----IREEDF	215
DB	208	vpdlipgtnyrciqdarnmidrlirelmgerrdsgeftldmlylmkqgnrlyptdee	267
QY	216	LDISIS-----MEDEHAAIRAKKGDGELLMNEDYOK	247
DB	268	rdqavtlllysgyetvstsmmaikylnhdpkalgelraehlafrerkrqdepljgdedvks	327

Tue Jul 30 08:40:06 2002

us-09-995-917a-1.rag

Page 23

[illegible]

Search completed: July 29, 2002, 13:59:58
Job time: 145 sec

Search completed: July 29, 2002, 13:59:58
Job time: 145 sec

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 21:27:14 ; Search time 217.88 Seconds
(without alignments)
9054.226 Million cell updates/sec

Title: US-09-995-917A-3

Perfect score: 1149

Sequence: 1 atggatgagccttcattg.....caaatcctgaagattag 1149

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_032802:*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107.6	9.4	1667	21	AAC35582
2	97.8	8.5	1608	18	AAT85306
3	97.8	8.5	1646	21	AAC48157
4	97.8	8.5	1649	21	AAC47986
5	96.8	8.4	1398	21	AAC48714
6	83	7.2	1461	22	AAS05196
7	82.4	7.2	6888	21	AAT85307
8	80	7.0	4937	18	AAT85307
9	71.2	6.2	1685	21	AAC39029

10	67.8	5.9	940	21	AAC37912	Arabidopsis thaliana
11	67.8	5.9	1673	21	AAC39858	Arabidopsis thaliana
12	59.6	5.2	1648	21	AAC42516	Arabidopsis thaliana
13	59	5.1	1503	22	AAS05171	Taxus cuspidata ox
14	52.4	4.6	660	21	AAC40136	Arabidopsis thaliana
15	50.8	4.4	1494	22	AAS05174	Taxus cuspidata ox
16	50.8	4.4	1530	22	AAS05177	Taxus cuspidata ox
17	49.8	4.3	1476	22	AAS05170	Taxus cuspidata ox
18	48.2	4.3	1482	22	AAS05193	Taxus cuspidata ox
19	48.6	4.2	1503	22	AAS05169	Taxus cuspidata ox
20	47.8	4.2	503	21	AAC56633	Eucalyptus grandis
21	44.4	3.9	1458	22	AAS05192	Taxus cuspidata ox
22	44.2	3.8	1491	22	AAS05194	Taxus cuspidata ox
23	44.2	3.8	1497	22	AAS05195	Taxus cuspidata ox
24	44.2	3.8	1506	22	AAS05168	Taxus cuspidata ox
25	43.8	3.8	1811	21	AAAV5431	DNA encoding novel
26	43.4	3.8	3927	21	AAAV70101	Plasmodium falciparum
27	43.4	3.8	4590	22	AAH24065	Yeast AOD9604-asso
28	43.4	3.8	9181	20	AAH43323	Streptococcus faecalis
29	41.2	3.6	1455	22	AAS05167	Taxus cuspidata ox
30	41.2	3.6	1524	22	AAS05175	Taxus cuspidata ox
31	41.2	3.6	2670	23	AAS94351	DNA encoding novel
32	40.8	3.6	2311	20	AAH13205	Enterococcus faecalis
33	40.6	3.5	1452	22	AAS05172	Taxus cuspidata ox
34	40	3.5	7498	24	ABL32257	Human immune system
35	39.6	3.4	1512	22	AAS05173	Taxus cuspidata ox
36	39.2	3.4	9810	24	ABL32426	Human immune system
37	38.8	3.4	2760	16	AAO85425	Tumour suppressor
38	38.6	3.4	6031	22	AAH46622	S. epidermidis open
39	38.6	3.4	1287	22	AAH53328	S. epidermidis open
40	38.6	3.4	3573	22	AAH54737	Pinus radiata trans
41	38	3.3	268	21	AAC56889	Human digestive system
42	38	3.3	463	22	AAK89102	Human digestive system
43	38	3.3	463	22	AAK89102	Human liver associated
44	38	3.3	463	22	AAK89102	Human liver associated
45	38	3.3	463	22	AAS31853	Human liver associated

ALIGNMENTS

RESULT 1	
AAC35582	standard; DNA; 1667 BP.
ID	AAC35582
XX	
AC	AAC35582;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 10705.
XX	
KW	Hybridisation assay; genetic mapping; gene expression control;
KW	protein identification; signal transduction pathway;
KW	metabolic pathway; promoter; termination sequence; ss.
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
ED	06-SEP-2000.
XX	
FE	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135529.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.

```

PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159337.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.
PR 21-OCT-1999: 99US-0160741.
PR 21-OCT-1999: 99US-0160767.
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161922.
PR 28-OCT-1999: 99US-0161933.
PR 29-OCT-1999: 99US-0162142.

```

Query Match 9.4%; Score 107.6; DB 21; Length 1667;

Best Local Similarity 56.1%; Pred. No. 5.9e-19; Matches 240; Conservative 0; Mismatches 164; Indels 24; Gaps 1;

```

QY 675 agaaccatgcagccattagagccagaaggagatgaggaaacctttgaattgaggagaattta 734
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1002 agaacatttgctataaggagagaacacgacctgacacgtctctcgacgatalat 1061
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 735 tcgagaagaatgaatttcataatggtgattctgaggcactagaatggtgtaataatcgt 794
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1062 taactcgaatgaattcctcgcagctgacgtctgacatcttgagacatccaagattgcaacgattgc 1121
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 795 caaagcgtacatagaagaactactatcgtatataatlcanaagaatgltgattccaaa 854
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1122 taatggtgctcttagaagaactactcagacttagaactccaacgcttatttaaccacaa 1181
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 855 ggggtggaagggtgttcccaatcttcacagcagatcattcgtacctctctcttcatagaaa 914
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1182 aggttggagaattctacgtatcacagaagagatlaactatgatactctctttagaaga 1241
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 915 tcctttgaatttaattcccatgagatggaaccaaaagca----- 952
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1242 tccaatgactcttaaccatgagatgagatggaagaagccttagaatccaagagctattt 1301
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 953 --cggcgcttggagagagaaaggtatgtcctggtggtgaaacttggcgaagctccaat 1010
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1302 ctactccttggaggtgaggttaggtcttggccctggaaagaaactagaatctcggaggt 1361
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1011 tgccttctctctatcatcttctgctcctctctataggtggaataatgaatcagaatgaat 1070
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1362 ctcaagctctctctactactcttgttaacaaatatagatggaaagaaatggaagaaacaa 1421
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1071 gccaatcg 1078
    |||||
DB 1422 attaatgg 1429

```

RESULT 2

AA785306
ID AA785306 standard; cDNA to mRNA; 1608 BP.
XX
AC AA785306;
XX
DT 14-APR-1998 (first entry)
XX
DE Arabidopsis thaliana cytochrome P450-type hydroxylase cDNA.
XX
KM Cytochrome P450-type hydroxylase; identification; brassinosteroid;

KM brassinosteroid inhibitor; modified plant; recombinant production;
KM teasterone; ds.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 48..1467
FT /tag= a
FT /product= cytochrome_P450-type_hydroxylase

PN W09J35986-A1.
XX
PD 02-OCT-1997.
XX
PD 27-MAR-1997; 97WO-EP01586.
XX
PF 27-MAR-1996; 96US-0622166.
XX
PR (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
PA Altman T, Koncz C, Mathur J, Szekeres MA;
XX WPI: 1997-489649/45.
DR P-PSDB: AAM27153.
XX

PT New isolated plant cytochrome P450-type hydroxylase gene - used to
PT identify substances acting as brassino-steroid(s) or brassinosteroid
PT inhibitors for the production of modified plants
XX
XX Claim 1: Pages 44-46; 77pp: English.

CC The present sequence encodes Arabidopsis thaliana cytochrome
CC P450-type hydroxylase. The hydroxylase can be used to identify
CC brassinosteroids or brassinosteroid inhibitors, useful to produce
CC plants with modified physiological and/or phenotypic
CC characteristics. The modified plants may show, e.g. stimulation of
CC growth, increased cell elongation, increased wood production,
CC accelerated seed germination at low temperatures, an increase in
CC dry weight, repressed anthocyanin production during growth in light
CC and/or inhibited de-etiolation which is induced, e.g. by cytokinin,
CC in the dark or an increase in stress tolerance. The hydroxylase or
CC its coding sequence can also be used for the recombinant production
CC of compounds, e.g. teasterone.

XX Sequence 1608 BP; 382 A; 374 C; 374 G; 478 T; 0 other;

Query Match 8.5%; Score 97.8; DB 18; Length 1608;
Best Local Similarity 58.9%; Pred. No. 2.5e-16;
Matches 168; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

```

QY 671 atgaagacatgcagccattagagccagaaggagatggggaactttgaattgggaag 730
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 961 aggaagagcatgaaagaattagagcaatggaaggtgttcgtatgacttggaatggaatg 1020
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 731 attatcagaagatggaattaccatcgaatgtgtttcttcttgaggactcgaatggttaata 790
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1021 attcaagatcaatggtccatcacacaacgtgtgttcaatgagacgctacgagtggtctaca 1080
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 791 tgcgcaagactgtacatagaaagctcatcatgatatataatccaagaatgatgtatc 850
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1081 tcaatcgcggtgtttttagacagtgcaatgacgagatgtttagatccaagaagtataaattc 1140
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 851 caaagggtggaagggtgttcccaatcttcacagcagatcattcgtacctctctcaatg 910
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1141 caaaagggtggaagatattcctcatcggtttagagcggtctatttagaccacaaacattca 1200
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 911 aaatcctttgaaattcaatccatgagatggaacccaagaacgag 955
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1201 aagatgtcgcaacttcaacccttggagatggaagagcaactcgg 1245

```

RESULT 3

AAC48157
ID AAC48157 standard; DNA; 1646 BP.
XX AAC48157;
AC
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 56467.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
FE 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137328.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145513.
PR 27-JUL-1999; 99US-0145518.
PR 27-JUL-1999; 99US-0145519.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.

```
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159350.
PR 14-OCT-1999; 99US-0159351.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 8.5%; Score 97.8; DB 21; Length 1646;
Best Local Similarity 58.9%; Pred. No. 2,6e-16;
Matches 168; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

```
QY 671 atgaagacatgcagccatagagacaaagaagatggagacttgaattgggaag 730
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 972 aggaagaagcctgaagaagattlaagggcaatgaagaagtgatcgtatagtctgaatggagtg 1031
QY 731 atatacagaagaatgaatcaactcaatgtgtatctcgaaggaactacgtatggttaata 790
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1032 attacaaatgaatgcattcacacaaatgtgtgttaatgagcgtaagtggtctaaca 1091
QY 791 tcgtcaagactgtacataagaagaagcctcatgatataatcaagaatattgtgattc 850
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1092 tcatcgcgcggtgttctccaaagctgtcaatgaagatgtgtgaatcacaagaatttaaatc 1151
QY 851 caaagggtggaagtggttcccaatctcacagcagttacatctgtaccctctcttcacgt 910
```

```
Db 1152 caaaaggctggaagatattctcatcgttttagagcgttcatctagaccacacttca 1211
QY 911 aaatccttctgaatttaaccatgagatgagacccaagaagagg 955
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1212 aagatgtctgcacttcaacccttgagatgagacagaacactcgg 1256

RESULT 4
AAC47986 standard; DNA; 1649 BP.
ID AAC47986.
AC xx
AC AAC47986.
DT 18-OCT-2000 (first entry)
XX xx
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 55845.
XX xx
KM Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX xx
OS Arabidopsis thaliana.
PN EP1033405-A2.
PD 06-SEP-2000.
XX xx
XX 25-FEB-2000; 2000EP-0301439.
FE
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
```

[illegible]

PR	12-AUG-1999;	99US-0148341
PR	13-AUG-1999;	99US-0148565
PR	13-AUG-1999;	99US-0148664
PR	16-AUG-1999;	99US-0149336
PR	17-AUG-1999;	99US-0149175
PR	18-AUG-1999;	99US-0149426
PR	20-AUG-1999;	99US-0149722
PR	20-AUG-1999;	99US-0149723
PR	20-AUG-1999;	99US-0149929
PR	23-AUG-1999;	99US-0149902
PR	23-AUG-1999;	99US-0149930
PR	25-AUG-1999;	99US-0150566
PR	26-AUG-1999;	99US-0150884
PR	27-AUG-1999;	99US-0151065
PR	27-AUG-1999;	99US-0151066
PR	30-AUG-1999;	99US-0151080
PR	30-AUG-1999;	99US-0151303
PR	31-AUG-1999;	99US-0151438
PR	01-SEP-1999;	99US-0151930
PR	07-SEP-1999;	99US-0152363
PR	10-SEP-1999;	99US-0153070
PR	13-SEP-1999;	99US-0153758
PR	15-SEP-1999;	99US-0154018
PR	16-SEP-1999;	99US-0154033
PR	20-SEP-1999;	99US-0154779
PR	22-SEP-1999;	99US-0155139
PR	23-SEP-1999;	99US-0155468
PR	24-SEP-1999;	99US-0155569
PR	28-SEP-1999;	99US-0156450
PR	29-SEP-1999;	99US-0156596
PR	04-OCT-1999;	99US-0157117
PR	05-OCT-1999;	99US-0157753
PR	06-OCT-1999;	99US-0157855
PR	07-OCT-1999;	99US-0158020
PR	12-OCT-1999;	99US-0158232
PR	12-OCT-1999;	99US-0158366
PR	13-OCT-1999;	99US-0159294
PR	13-OCT-1999;	99US-0159294
PR	14-OCT-1999;	99US-0159329
PR	14-OCT-1999;	99US-0159329
PR	14-OCT-1999;	99US-0159330
PR	14-OCT-1999;	99US-0159331
PR	14-OCT-1999;	99US-0159637
PR	14-OCT-1999;	99US-0159638
PR	18-OCT-1999;	99US-0159584
PR	21-OCT-1999;	99US-0160747
PR	21-OCT-1999;	99US-0160747
PR	21-OCT-1999;	99US-0160768
PR	21-OCT-1999;	99US-0160770
PR	21-OCT-1999;	99US-0160814
PR	21-OCT-1999;	99US-0160815
PR	22-OCT-1999;	99US-0160980
PR	22-OCT-1999;	99US-0160981
PR	22-OCT-1999;	99US-0160989
PR	22-OCT-1999;	99US-0161040
PR	22-OCT-1999;	99US-0161405
PR	23-OCT-1999;	99US-0161406
PR	26-OCT-1999;	99US-0161359
PR	26-OCT-1999;	99US-0161360
PR	26-OCT-1999;	99US-0161361
PR	28-OCT-1999;	99US-0161920
PR	28-OCT-1999;	99US-0161992
PR	29-OCT-1999;	99US-0161993
PR	29-OCT-1999;	99US-0162142

Query match	8.5%;	Score 97.8;	DB 21;	Length 1649;
Best local Similarity	58.9%;	Pred. No. 2.6e-16;		
Matches 168;	Conservative	0;	Mismatches 117;	Indels 0;
				Gaps 0;

QY 671 atgaagaacatgcagccattagagccaagaaggatcgggaaactttgaaattggagaag 730
| | | | | | | | | | | | | | | | | | |
Db 974 aggaagaagcatgaaaagattaaaggcaatgaaagtgtacttgatagtccttgaattgsagtg 1033


```

PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149802.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156586.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159684.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match      8.4%; Score 96.8; DB 21; Length 1398;
Best Local Similarity 53.0%; Pred. No. 4, 5e-16;
Matches 249; Conservative 0; Mismatches 197; Indels 24; Gaps 1.

QY 676 gaacatgcagccattagacgaagaagggaatggggaactttgaattgggaattat 735
   ||||| || || || || || || || || || || || || || || || || || || ||
Db 916 gagcattggcattccaggaagaagaacgacagacgaacccactcgtcttgagagctg 975

QY 736 cagaagatggaattcaactaatgtgattcttcgaagccactcagatggttaatcgctc 795
   || || || || || || || || || || || || || || || || || || || || || ||
Db 976 aagcattgaagttcactcagctcagctgattttagacalcagatcggcaacgactgatt 1035

QY 796 aagacgttacaatagaagaactactacatgatatataatcaagaatattgattccaaag 855
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1036 aatggggtccaaagaaactactcgtgacttggaaatcaactcaggttatttaaccocaa 1095

QY 856 ggggtggaagggtgttccaatcttcacagcagttacatcttgatccctctcattgaata 915
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1096 ggaatggaattattgtatatacagcaggaatlaattacgattcaatcttattgaagac 1155

QY 916 cctttggaattatccatgagatgagaccaaaagca----- 952
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1156 ccaatgattccttaatccatgagatgagatgagaagaagagcttgagatcacaaactatgc 1215

QY 953 -cgagcgttggagaggaagtaaggatgtcctcgtgtgtgtaacttggcgaagctccaaat 1011
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1216 ttgtgttttgaggtggagaaagccttgcctcgtgaagaaactagagattgcagatc 1275

QY 1012 gcttctcctcaatcaatcttgcctcctcctaagtgtggaataaagttaagtatgaatg 1071
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1276 tcgagcttccatcattcacttgtaacgatacagatggaagaataaggaagggtatgaa 1335

QY 1072 ccaatgcgcaccccttaacgctgaggttaagaagaagagcagcttggagat 1121
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1336 ttaatggtgttccgagagtttttgacacaaaaggtccatcttgat 1385

RESULT 6
AAS05196
ID AAS05196 standard; DNA; 1461 BP.
XX
AC AAS05196;
XX
DT 07-SEP-2001 (first entry)
XX
DE Taxus cuspidata oxygenase enzyme DNA #19.
XX
KW Oxygenase; Japanese yew; Taxol; taxoid; ds; Taxol biosynthetic pathway;
transgenic organism.
XX
OS Taxus cuspidata.
XX
PN W0200134780-A2.
XX
PD 17-MAY-2001.
XX
PF 13-NOV-2000; 2000WO-US31254.
XX
PR 12-NOV-1999; 99US-0165250.
XX
PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX
PI Croteau RB, Schoendorf A, Jennewein S;
XX
DR WPI: 2001-355489/37.
XX
P-PSDB: AAU02839.
XX
PT Novel nucleic acid and amino acid sequences, isolated from the Taxus
taxoids, intermediates within the Taxol biosynthetic pathway, and
other taxoid derivatives -
XX

```


PS Claim 4; Page 130-131; 141pp; English.

CC The sequence represents a DNA which encodes a *Taxus cuspidata* oxygenase.
 CC The sequences isolated from the *Taxus* genus, and the respective
 CC oxygenases are useful for the synthetic production of taxol and related
 CC taxoids, as well as intermediates within the Taxol biosynthetic pathway,
 CC and other taxoid derivatives. The sequences also can be used to make
 CC transgenic organisms that either produce the oxygenases for subsequent *in*
 CC vitro use, or produce the oxygenases *in vivo* so as to alter the level of
 CC taxol and taxoid production within the transgenic organism. The oxygenase
 CC nucleic acids and amino acids are useful for isolating the polynucleotide
 CC and polypeptide sequences corresponding to full-length oxygenases.

XX Sequence 1461 BP; 412 A; 283 C; 361 G; 405 T; 0 other;

Query Match 7.2%; Score 83; DB 22; Length 1461;
 Best Local Similarity 48.1%; Pred. No. 2.4e-12;
 Matches 298; Conservative 0; Mismatches 315; Indels 6; Gaps 2;

OY 4 ggaatgacctcatttggaagaactatctctctcacaacctcacaatagatcagatccatc 63
 DB 154 ggaatgacctcatttggaagaactatctctctcacaacctcacaatagatcagatccatc 213
 OY 64 ggtacatctctgcacacacgtgttcacggtatggaagaagtgtcaaatatataatgt 123
 DB 214 cgaatgacctcatttggaagaactatctctcacaacctcacaatagatcagatccatc 273
 OY 124 ggtggaaga--ggaatgacctcatttggaagaactatctctcacaacctcacaatagatcagatccatc 180
 DB 274 ggaatgacctcatttggaagaactatctctcacaacctcacaatagatcagatccatc 333
 OY 181 gaaagggaagtgtttacatcgatcgaatcacaagaagcgaatgacatcgcgcgacaat 240
 DB 334 gaaagggaagtgtttacatcgatcgaatcacaagaagcgaatgacatcgcgcgacaat 393
 OY 241 tccctctatctagcacaacggaatcacaagaagaataaataatgtatattatagctc 300
 DB 394 ggaatgacctcatttggaagaactatctctcacaacctcacaatagatcagatccatc 453
 OY 301 atcaatctcacaagaactggaacactgactctctcacaacctcacaatagatcagatccatc 360
 DB 454 tgaagatggaagcgcacactgactctcacaacctcacaatagatcagatccatc 513
 OY 361 atcctaaagtcatacgaagaatcgcgaagaagtcgaatcacaagaagaatgaatgttt 420
 DB 514 ggaatggaagaatcgcgaagaagtcgaatcacaagaagaatgaatgttt 573
 OY 421 acctcgaagtgttatgtaaaccaactcttgacatcagaagcagaagaaccagacatc 480
 DB 574 gttctgaaacttgatgcgaagaagatctgacttaccctcacaaga--agaatggga 630
 OY 481 tatgatacgaagaattttatctatatacgaagaaggtttatcctcacaagaaccgctt 540
 DB 631 catattataaagcttcgacgatttcgtaggagcgtgctcctcctccccaataatc 690
 OY 541 ccaaggaaaggtatcacaacacgaatgaaggttagatccacatgtaatacacaac 600
 DB 691 cctggaaacacattatcgagaagaattcggcgaagggaattctgttaaaagaatcacc 750
 OY 601 gcaattatagagaatga 619
 DB 751 aagtgtataaagaaga 769

RESULT 7

AAAS9599
 ID AAAS9599 standard; DNA: 6888 BP.

AC AAAS9599;

DT 14-NOV-2000 (first entry)

XX

DE DNA encoding a cytochrome P450 enzyme designated DMF4.

XX DMF4: cytochrome P450 enzyme; brassinosteroid; 22alpha-hydroxylation;
 KW plant phenotype; cell elongation; ss.

XX Arabidopsis sp.

XX Key Location/Qualifiers

FT promoter

FT TATA-signal

FT CDS

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

The present sequence encodes a DMF4 polypeptide. The polypeptide is a
 cytochrome P450 enzyme that mediates multiple steps in synthesis of
 brassinosteroids. Specifically, it mediates multiple

RESULT 9
AAC39029
ID AAC39029 standard; DNA; 1685 BP.
XX
AC AAC39029;
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 23101.
DE
XX Hybridisation assay; genetic mapping; gene expression control;
KW Protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140033.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147197.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 6.2%; Score 71.2; DB 21; Length 1685;
Best Local Similarity 49.4%; Pred. No. 3.8e-09;
Matches 230; Conservative 0; Mismatches 218; Indels 18; Gaps 1;

QY 636 aagaatttcgtgattcgaatttcgaatggaatggaacatgcagccattgagc 695
DB 1002 acatgtatcgaaacttcctagaagctgttaagctcgagcaaaagccataataaaga 1061
QY 696 caagaagaaggatggygaacttttgattggygaagattacgaagaatggaattcactca 755
DB 1062 aaacagtagaggaagaacacttaacatgtagagacaacgaggaatatgcccattgacaca 1121
QY 756 atgttgatttcgaggaactacgatgtgttaatatctgcacaagacgtcatatagaagaac 815

DB 1122 taagttatagtgaagcttgagatggycaagcatcatatccttcacatcagaagaac 1181
QY 816 tactcatgtataatttaagaagaatgatgtatccaaagggtggaagtgttccaat 875
DB 1182 atgtgttgattgtgaataaaggatattgttaacctaaaggatggaagtgtatgccact 1241
QY 876 ctcaacagcagctacatcctgtgacctctctctcatganaatcctttggaattcaatccat 935
DB 1242 gtttcggaatatcatcacaatccgaatatattttcaaacctcgaggtttcagccacc 1301
QY 936 gagatg-----gaccaaacgacggtgttggaaggagtaaggt 977
DB 1302 tagattcgaggtaaatccgaagccgaatacatcatcctgttggaagtgtgattcatc 1361
QY 978 atgtcctgtgtgtaactctgcaagctccaattgtcttctcctcatcactctgtcct 1037
DB 1362 tgtcccggaacgaactctgcgaagttcaaatcttataattcttcaccattagttcc 1421
QY 1038 ctccctaaagtgaanaataaagtcagatgaatgccaatcgcgac 1083
DB 1422 caattccgatggaagtgaaggaggaagaagaataacagttac 1467
RESULT 10
AAC37912
ID AAC37912 standard; DNA; 940' BP.
XX
AC AAC37912;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 19104.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144881.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145210.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148568.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155148.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.


```

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145102.
PR 23-JUL-1999; 99US-0145114.
PR 23-JUL-1999; 99US-0145212.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145291.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 30-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

```

```

PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160982.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.9%; Score 67.8; DB 21; Length 1673;
Best Local Similarity 54.0%; Pred. No. 3.1e-08;
Matches 181; Conservative 0; Mismatches 127; Indels 27; Gaps 1;

QY 727 gaagattatcagaagatgaattcaactcaatgtgtgattcttgaggcactcagatgtgt 786
DB 1092 gaattatagacacacattcaccacacatggtttagcaatggtgtcttcgctagca 1151
QY 787 aatatacgtcaagactgtcatagaaagctactcatatgatatctaataacgaagatgtg 846
DB 1152 aacagaccccttgtgttcttcgaaggcgtgcagatgtcttgagatcaagatatca 1211
QY 847 attcacaagggtgtggaaggtgttccaaacttcacagcagatcactgtgacctctct 906
DB 1212 attccgctgtgttgattgtgtggtgtgcacacttcacgcgttacttcttgatccctgaatc 1271
QY 907 catgaatacctcttgaattcaatcccatgagatg----- 942
DB 1272 tatgagaacccattgtatcattcaatcactgagatgtggaagaaatgatttgggga 1331
QY 943 --accacaagcagcgcttggaggaagtaaggatgtcctgtgtgtgaactgtgc 999
DB 1332 tctaaacgttcaatgtggttggatgatgagattgagacttgtgtgagtgacagagtttcg 1391
QY 1000 aagctcacaatgtcttcttcctcatcatcattgt 1034
DB 1392 cggctcaaatgtcaatcttccctcatcatctgt 1426

RESULT 12
AAC42516
ID AAC42516 standard; DNA; 1648 BP.
XX AAC42516;
AC
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35853.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.

```

PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 23-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134770.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136621.
PR 27-MAY-1999; 9905-0136992.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.

PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144332.
PR 20-JUL-1999; 9905-0144632.
PR 21-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0145086.
PR 22-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 05-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 10-AUG-1999; 9905-0147935.
PR 11-AUG-1999; 9905-0148171.
PR 12-AUG-1999; 9905-0148319.
PR 13-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 16-AUG-1999; 9905-0148684.
PR 17-AUG-1999; 9905-0149368.
PR 18-AUG-1999; 9905-0149175.
PR 20-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 23-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150930.
PR 26-AUG-1999; 9905-0150966.
PR 27-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151065.
PR 30-AUG-1999; 9905-0151080.
PR 31-AUG-1999; 9905-0151303.
PR 01-SEP-1999; 9905-0151438.
PR 07-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.

PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158322.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	28-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	29-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

[illegible]

XX AC AAS05171:
XX XX
XX DT 07-SEP-2001 (first entry)
XX DE Taxus cuspidata oxygenase enzyme DNA #6.
XX KM Oxygenase; Japanese yew; Taxol; taxoid; ds; Taxol biosynthetic pathway;
XX KM transgenic organism.
XX OS Taxus cuspidata.
XX PN W0200134780-A2.
XX PD 17-MAY-2001.
XX PF 13-NOV-2000; 2000MO-US31254.
XX PR 12-NOV-1999; 99US-0165250.
XX (UNITW) UNIV WASHINGTON STATE RES FOUND.
XX PA Croteau RE, Schoendorf A, Jennewein S;
XX PI WPI: 2001-355489/37.
XX DR P-PSDB: AAU02826.
XX PT Novel nucleic acid and amino acid sequences, isolated from the Taxus
XX PT genus, useful for the synthetic production of Taxol and related
XX PT taxoids, intermediates within the Taxol biosynthetic pathway, and
XX PT other taxoid derivatives -
XX PS
XX PS Claim 4; Page 99-100; 141pp; English.
XX CC
XX CC The sequence represents a DNA which encodes a Taxus cuspidata oxygenase.
XX CC The sequences isolated from the Taxus genus, and the respective
XX CC oxygenases are useful for the synthetic production of Taxol and related
XX CC taxoids, as well as intermediates within the Taxol biosynthetic pathway,
XX CC and other taxoid derivatives. The sequences also can be used to make
XX CC transgenic organisms that either produce the oxygenases for subsequent in
XX CC vitro use, or produce the oxygenases in vivo so as to alter the level of
XX CC Taxol and taxoid production within the transgenic organism. The oxygenase
XX CC nucleic acids and amino acids are useful for isolating the polynucleotide
XX CC and polypeptide sequences corresponding to full-length oxygenases.
SQ Sequence 1503 BP; 406 A; 331 C; 324 G; 441 T; 1 other;

Query Match 5.1%; Score 59; DB 22; Length 1503;
Best Local Similarity 45.0%; Pred No. 7.1e-06;
Matches 264; Conservative 0; Mismatches 320; Indels 3; Gaps 1

QY 2 tgggatggccttcattggagaacatattcttcttcacaaacctcatagatcagactcca 61
DB 176 tagcttccttccttcattgggagacatcacattcttgaggcgacttcgacatcagaacac 235
QY 62 tcggagacattcttgacaacacgtgttccacggtataggaaaagtgttcaagtcaaatatc 121
DB 236 ctcaagacgtcttcgtatgtaggaggtgaagaatactcggtgttatctcaagactcgatag 295
QY 122 gtggaggaaaagcagtagtctcatcttgaccgaacatcacatgtgtcatactccaagc 181
DB 296 ttgggacatcccccagttgtactctcgtcgggccgtgagggaacacgccttctctcacaag 355
QY 182 aagggaagtgtttcatcgcgatcttcacaaaagcgtatcatatcaactctcggaataat 241
DB 356 aggaacaactcgttgaggcgctatcgtcccaaatcttcgcgaagaacataattggaaatat 415
QY 242 cccctcatttggccacggagaaatccacaggaacataaaaattttattatagcttca 301
DB 416 ccattctgtccaaaagagggaaggacatcgtcatatttcagtcgcacttgcgcgtctt 475
QY 302 tcaatctcaaaaagtcgaacctgaattcttcaactgctcagagaagaacctctatctcga 361


```

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147312.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155386.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

```

```

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 4.6%; Score 52.4; DB 21; Length 660;
Best Local Similarity 48.1%; Pred. No. 0.00032;
Matches 149; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

```

```

QY 1 atggatgagccttcatttgagaactattcttctcaaacctatagatcagctcc 60
DB 224 atggattaccgatactcgagagacatgcactctcttgagcccatgatatccg 283
QY 61 atcggatcattcttgcaacaacgctgttcaagytatgaaagtgatcaatcata 120
DB 284 atcccaaccttgcaagaagatgataagtgacggccattgttcgacaacatc 343
QY 121 tctgtgtaaaagcaatgactcatgtgacccaagaactacacgttcactccaac 180
DB 344 ttcggatccaaacacgctgttctgacagaagcgtatcatatcttgaggtttcggcaa 403
QY 181 gaagggaagtgtttacatcgatataccaagaagcgtatgacatcttcgcaaatat 240
DB 404 gagaacaagcttcttgactagctatccagagccattccaagcatttgaaagaa 463
QY 241 tccctctatagccacgagagaatccaggaactaaagaatgtatttgcctc 300
DB 464 acgtgttccctaaacatgaaacatccacaagaacgltcaacaatcaglttcaactc 523
QY 301 atcaatctca 310
DB 524 ttgctctga 533

```

Result 15

ID AAS05174 standard; DNA; 1494 BP.

AAS05174

AC AAS05174;

DT 07-SEP-2001 (first entry)

DE Taxus cuspidata oxygenase enzyme DNA #9.

KW Oxygenase; Japanese yew; Taxol; taxoid; ds; Taxol biosynthetic pathway;

KW transgenic organism.

OS Taxus cuspidata.

PN WO200134780-A2.

PD 17-MAY-2001.

PF 13-NOV-2000; 2000WO-US31254.

XX 12-NOV-1999; 99US-0165250.

XX (UNITIV) UNIV WASHINGTON STATE RES FOUND.

XX Croteau RB, Schoendorf A, Jennewein S;

XX WPI: 2001-355489/37.

XX P-PSDB; MAU02829.

XX Novel nucleic acid and amino acid sequences, isolated from the Taxus

XX genus, useful for the synthetic production of Taxol and related

XX taxoids, intermediates within the Taxol biosynthetic pathway, and

XX other taxoid derivatives -

XX Claim 4; Page 101; 141pp; English.

XX The sequence represents a DNA which encodes a Taxus cuspidata oxygenase.

XX The sequences isolated from the Taxus genus, and the respective

XX oxygenases are useful for the synthetic production of Taxol and related

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 21:23:04 ; Search time 1717.1 Seconds
(without alignments)
9031.507 Million cell updates/sec

Title: US-09-995-917A-3

Perfect score: 1149
Sequence: 1 atggagatgccttcattg.....caaatccttggaattag 1149

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estchum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query length	DB ID	Description
1	233.4	20.2	755 10	BI926091 EST545980
2	206	17.9	574 9	AM61836 EST325334
3	204.4	17.8	574 9	AM615967 EST325333
4	181.8	15.8	615 9	AW398669 EST309169
5	164.6	14.3	609 10	BI204653 EST522693
6	163	14.2	678 10	BI204446 EST522486
7	157	13.7	533 10	BE450142 EST401029
8	152.2	13.2	667 10	BF324723
9	143.4	12.3	621 10	BI204438 EST522478
-10	128.6	11.2	522 10	BI204438 EST522478
11	127.2	11.1	573 10	BI206239 EST524279
12	124.8	10.9	660 10	BI266196
13	120.2	10.5	582 10	BI205718
14	118.2	10.3	642 10	BE315722
15	116	10.1	533 10	BF050501
16	110.4	9.6	477 10	BF052296
17	108	9.4	543 10	BI432565

18	98.4	8.6	586 10	BM409833	BM409833	EST584160
19	96.8	8.4	299 10	BI072383	BI072383	C074P580
20	95	8.3	574 10	BI787249	BI787249	98973B09.
21	93	8.1	312 10	D15214	D15214	R1CC0279A.R
22	92.8	8.1	594 10	BI928055	BI928055	EST547944
23	91.4	8.0	536 9	A1994171	A1994171	701499941
24	90.4	7.9	524 9	AM690507	AM690507	NF035D01S
25	89.2	7.8	347 10	BI071486	BI071486	C058P32U
26	87.8	7.6	432 10	Z26124	Z26124	ATTS1396.Gr
27	87.2	7.5	535 10	BI432249	BI432249	EST53501B
28	86.6	7.5	788 10	BI970911	BI970911	GM830011B
29	86.4	7.5	684 10	Z29017	Z29017	ATTS2030.Gr
30	85.4	7.5	1023 9	A1943419	A1943419	MF02A4.MF
31	83.8	7.3	343 10	T22324	T22324	4332.Lambda
32	83.6	7.3	608 9	BE124630	BE124630	EST393665
33	82.8	7.2	658 10	BI263183	BI263183	NF08B09P
34	82.2	7.2	360 10	T43151	T43151	6414.Lambda
35	82	7.1	501 10	T43286	T43286	6549.Lambda
36	81.8	7.1	619 10	BF424875	BF424875	SUS2B05.Y
37	81.6	7.1	760 12	BM479918	BM479918	BOGJT72TF
38	81.2	7.1	505 9	AM201483	AM201483	9703f10.Y
39	80.8	7.0	461 9	AM759817	AM759817	s154d10.Y
40	80.8	7.0	508 10	BE595940	BE595940	8U67g12.Y
41	80.4	7.0	286 10	Z17988	Z17988	ATTS0472.AC
42	80.4	7.0	469 9	A1993171	A1993171	701495585
43	80	7.0	686 9	BE039087	BE039087	AB09E01.A
44	80	7.0	339 10	T22325	T22325	4333.Lambda
45	79.6	6.9				

ALIGNMENTS

RESULT 1
BI926091
LOCUS
DEFINITION
EST545980 tomato flower, buds 0-3 mm mRNA linear EST 18-Oct-2001
clone CTOA28L21 5' end, mRNA sequence.
BI926091
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Lycopersicon
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 755)
van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
Uterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
Generation of ESTs from tomato flower tissue, 0-3 mm buds (2001)
Unpublished (2001)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
location/Qualifiers
1..755
/organism="Lycopersicon esculentum"
/cultivar="RA496"
/db_xref="taxon:4081"
/clone="CTOA28L21"
/clone_lib="tomato flower, buds 0-3 mm"
/tissue_type="flower"
/dev_stage="0-3mm buds"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers

were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

BASE COUNT
ORIGIN

226 a 140 c 124 g 265 t

Query Match 20.2%; Score 232.4; DB 10; Length 755;
Best Local Similarity 63.0%; Pred. No. 6.5e-46;
Matches 359; Conservative 0; Mismatches 211; Indels 0; Gaps 0;

QY 2 tggagtgagcttcattgagagaactattcttctccaacatcatagatcaactcoa 61
DB 185 ttggatggccttacttggtggaaccccttcttttgagacctatccttttaattca 244
QY 62 tcgfatcattcttgacaacagtglttcaaglatgtaaaagtgltcaagtaatat 121
DB 245 ttggatcttcttcttcaacaaatggtttagtggaagtggtcaagtcacatttat 304
QY 122 gtggtggaagacagtagctcatgttgaccaactcaactgttctactcaaacg 181
DB 305 tttttcccccacagtggtgcatgacacacaccttaattactcatattacaaacg 364
QY 182 aaggggaagtggttatactcgatlatccaaagcagatgacatctcgcgaatat 241
DB 365 aagatgaattatttgcgtgtagttaccaaagcattcattgatacttgcaaaagttt 424
QY 242 ccccttattagccaccgagagaatcacaagaactaaatgttattatagcttca 301
DB 425 catgcttctggtggttgccgacacacatmaaaagccttggaattggtcattatccta 484
QY 302 taatctcaacaaagtcgaacactgacttcttcaactgcgcagaagaacctctatcga 361
DB 485 tcgacacacatttgaagtcgaacctgatttattatggtgaaacattatgacattcaga 544
QY 362 tactaaagcatggaanaatttcggaagtcgaattccataaagaagtaaaagtta 421
DB 545 ttgtccaatcattggaanaagatmaaacatcagatcagtggaagacgaagaaagtttt 604
QY 422 ctctcaagtgatgtgaacaacactcttgagctcaagcagaagccagaagactt 481
DB 605 catctaatgagatggaacgaacagtgatcttgaatttaactccagatmaacaaagtcgat 664
QY 482 atgattgcaagatttttataatagaagaggttatactcttaccataaccccttc 541
DB 665 taattcttcaagattttcttctttatagagatttaattctttacattatcattac 724
QY 542 caggaaacggttatcaaacacgaattaag 571
DB 725 ctggaacatctcatatgcmaagacagtcagcag 754

RESULT 2
AM615836 574 bp mRNA linear EST 18-MAY-2001
LOCUS EST325334 tomato flower buds 0-3 mm, Cornell University
DEFINITION Lycopersicon esculentum cDNA clone cTOA17E20 5', mRNA sequence.
ACCESSION AM615836
VERSION AM615836.1 GI:7321721
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 574)
REFERENCE
AUTHORS van der Hoeven, R.S., Bezzeredes, J.L., Matern, A.L., Holt, I.E., Liang
F., Hansen, T., Craven, M.B., Bowman, C.L., Ronning, C.M., Niernan, W.,
Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
TITLE Generation of ESTs from tomato flower tissue, 0-3 mm buds
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/Index.html>
5 prime sequence.

FEATURES
source
Location/Qualifiers

1..574

/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOA17E20"
/clone_1db="tomato flower buds 0-3 mm, Cornell University"

/tissue_type="flower"
/dev_stage="0-3mm buds"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Tanksley; Flower buds and flowers were
taken from greenhouse plants (4-8 wks old, TA496). They
were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."
1 others

BASE COUNT 194 a 96 c 108 g 175 t

Query Match 17.9%; Score 206; DB 9; Length 574;
Best Local Similarity 61.3%; Pred. No. 1.6e-39;
Matches 332; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY 75 gcaacaagctgtttcacggtatggaagtggttcagtcacaaatataigtgtggaagaac 134
DB 1 gcacacacatttctttagttagtggaagtggtcgaacacacatttatttttcccccaac 60
QY 135 agtagtctcatgtgacaaagaactcaacatgttcatcacttcaaaacgaaggaagttt 194
DB 61 agtggtgcatgtgacacaaagaccttaattactctcatatttcaaaacgaagatgaatt 120
QY 195 tacatcgatttccaaaagcagatgacatcttcgcgcaaatatccctctattagc 254
DB 121 tcagtgatcttattccaaagcccaatttcatgcttacttgccaaagtttcttgctg 180
QY 255 caccgagaatctcagaagaactaaanaatttattattatagcttcatcattcaacaa 314
DB 181 ttgttgccgacacacatmaaaagccttagaagtttcattatcattatcagacacattaa 240
QY 315 gtccgaacctgacttcttctcaactgcgagagaacctctctatctcgatactaagttc 374
DB 241 gtcctaaacctgagtttatttgaatgattgaacatttagacacttcagatttccaaatc 300
QY 375 gaaaattgcccgaagtgagatccataagaagaattaaatgttactcagttgat 434
DB 301 gaaagatmaaacattcaagatcagatctggaaggaagcgaagaagtttccatgaatgat 360
QY 435 ggtaaacgaactcttgagatcaagcagaagaagccagaagaacttatgtatgcaaga 494
DB 361 agtgaaagcaagtgacttggaatttactccagatattccacaaagtgacatttcttcaga 420
QY 495 tttttatcttataagaaggtttatctcttaccataacgcttcagaagcaggttga 554
DB 421 ttttcttcttattatgaagagatttatttctttaccattatcattacctcgaaactcatt 480
QY 555 tacaacgcaattaaagtgtagatccaatcgtaatatatacaaaacgcaattatagaaga 614
DB 481 tcgaagagcagtcgacagcttagaagtagaattttcgaactatcgaagcaattatagaaga 540
QY 615 ca 616
DB 541 aa 542

RESULT 3
AM615967 574 bp mRNA linear EST 18-MAY-2001
LOCUS AM615967
DEFINITION EST325333 tomato flower buds 0-3 mm, Cornell University
Lycopersicon esculentum cDNA clone cTOA17E16 5', mRNA sequence.

```

ACCESSION   AM615967
VERSION     AM615967.1
KEYWORDS    GI:7321720
SOURCE      tomato.
ORGANISM    Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1 (bases 1 to 574)
AUTHORS    van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E., Liang
            ,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W.,
            Frazer,C.M., Martin,G.B., Giovannoni,J.J., and Tanksley,S.D.
TITLE       Generation of ESTs from tomato flower tissue, 0-3 mm buds
JOURNAL     Unpublished (1999)
COMMENT    Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.

FEATURES
source      Location/Qualifiers
            1..574
            /organism="Lycopersicon esculentum"
            /cultivar="TA496"
            /db_xref="taxon:4081"
            /clone="CTOA17E16"
            /clone_1lb="tomato flower buds 0-3 mm, Cornell University"
            /tissue_type="flower"
            /dev_stage="0-3mm buds"
            /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
            XhoI; supplier: Tanksley; flower buds and flowers were
            taken from greenhouse plants (4-8 wks old, TA496). They
            were immediately frozen in liquid nitrogen and then
            size-separated while remaining frozen."

BASE COUNT  196 a          96 c          109 g          173 t

ORIGIN
Query Match      17.8%; Score 204.4; DB 9; Length 574;
Best Local Similarity 61.1%; Pred. No. 3.9e-39;
Matches 331; Conservative 0; Mismatches 211; Indels 0; Gaps 0;

QY 75 gcaacaagtglttcacggtatggaagtgltcaagtcataatatagtgtggaagc 134
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GCACACACATTTGTTAGTATGGAAAGTTCACAGTCACTTATTTTCCCCCAAC 60

QY 135 agtaagctcatgtgacgaagaactcaactgttcaacttcaaaaggaagtgct 194
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 AGTGTCATGTGACCAAGACCTTATTAATTCACTATTACAAAGAAATAGTTATT 120

QY 195 tacatcgaatacacaagaacgcatgacattctcgcaaaatattcccttattagc 254
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 TCAGTGTACTTATCCAAACCAATTCATGTTACTGTGCAAAAGTTTCCTGTGCGC 180

QY 255 caccggaagaattcacaggaagaactaaatgttattatagcttcaatcacaac 314
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 TGTTCGGCGACACATTAAGAGCTTAGAATGTTTCATTATCAGTATCAGCACCATTTAA 240

QY 315 gtcgaagaactgactctcactcgcgaggaacccctctcactcgaattcaagtcag 374
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GCTTAACCTGAGTTTATTAATGATGTTTCAACATTAGCAGCTTCAGATTCTCCAAATCAG 300

QY 375 gaaanaatgacgagaagatcgaaatccataaagaaglttaaatgttactcagtgat 434
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 GAAAGATTAACATCAAGTCAGATACGAGAGAGGCAAGCAAGTTTTCATTCGAAGTGAT 360

QY 435 ggttaaacacactctgagatcaagccaaagaacccaaagaagacttatgtatcgaaga 494
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 AGTGAAGCAAGTACTTGGATTAACTCCAGATATATCCAAAGTGCAATTAACTTTCACAA 420

QY 495 ttcttatctatacgaagggcttataccttacaataacgcttcaggaagggata 554
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 421 TTTTCTGCTTTATAGAGAGATTAACTTTTACATTATACATACCTGGAATCCATA 480
QY 555 tacaaagcaatgaaggttagatccaaatcgtaatacatcaaaacgcaatlaagaaga 614
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 TGCAGAGCAGCTCAGCAGGTAGAAATATCTTCACATATCAAGCAATATATAGAGA 540
QY 615 ca 616
Db 541 AA 542

RESULT      4
LOCUS       AM398669
DEFINITION EST309169 L. pennellii trichome, Cornell University Lycopersicon
ACCESSION   AM398669
VERSION     AM398669
KEYWORDS    AM398669.1 GI:6917139
SOURCE      EST.
ORGANISM    Lycopersicon pennellii.
            Lycopersicon pennellii.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1 (bases 1 to 615)
AUTHORS    Alcalá,J., Vrebalov,J., White,R., Matern,A.L., Lakey,J., Holt,I.E.,
            Liang,F., Hansen,T.S., Upton,J., Ronning,C.M., Craven,M.B., Fujii
            ,C.Y., Bowman,C.L., Nierman,W., Frazer,C.M., Venter,J.C., Martin
            ,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE       Generation of ESTs from wild tomato (Lycopersicon pennellii)
            trichomes
JOURNAL     Unpublished (1999)
COMMENT    Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.

FEATURES
source      Location/Qualifiers
            1..615
            /organism="Lycopersicon pennellii"
            /db_xref="taxon:28526"
            /clone="CLP74K2"
            /clone_1lb="L. pennellii trichome, Cornell University"
            /tissue_type="trichome"
            /dev_stage="mixed stages"
            /lab_host="SOLR"
            /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
            XhoI; Leaves of various stages were shaken in liquid
            nitrogen, shearing off trichomes. This procedure yielded a
            mixture of cells highly enriched for trichomes, with minor
            contamination by other types of leaf cells."

BASE COUNT  180 a          111 c          100 g          224 t

ORIGIN
Query Match      15.8%; Score 181.8; DB 9; Length 615;
Best Local Similarity 63.7%; Pred. No. 1.2e-33;
Matches 276; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 2 tggagatgaccttcattggaagaactattcttctcaaaacctcatgatacagctcca 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 TTGGATGGCCTTACTTGTGTAACCTTCTTTTGAAGCCTATCTCTTAATTCTA 241

QY 62 tccgtaacctcttgacaacaagtglttcaaggtatgaaagvglttcaagtcataatcat 121
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 TTGGTACTTTCCTTACAAACATTTGTTAGTATGGAAAGTGTCAAGTCACATTTAT 301

QY 122 gtggtgaaagaacgtagtctcatgtagcaagaactcaaacgttcatctactcaaaagc 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 302 TTTTCTCCCAACAGTGGTGTATGTGACCAAGACCTTAATTACTTATATACAAATG 361

```

QY	182	aagggaagttgtttccalcggaattccaaaagcagatgataccttctcgcaaat	241
QY	183		242
Db	362	AAGATTAAGTATTATTAGTGTAGTATTATCCAAACCAATTCATGTACTGTGGCAAGTTT	421
QY	242	cccttctatagccaccggaagaattccaggaaaactaaaatgttattatgaacttca	301
Db	422	CATGCTGTGTGCTGTGTGGTGCACACACATAAAGCTTAGAATGTTTCATTATACATTA	481
QY	302	tcaactacaaagtcgaaacctgacttcttcaactgcgcaagaacctctatctga	361
QY	482	TCAGACACCTTTAAGTCTAAACCTGGATTATTAAATGATCTTTAAACATTAGCACTTCACA	541
QY	362	tactaagtcactgagaaaatitgcgagaagtcgaaactcatataagaagttaaatgttta	421
Db	542	TCTCTTCATTCATGAGAAAGATTAACATCAAGTACGTACTGGGAGAGGCAAGAAATTTT	601
QY	422	ccttcagtgatat	434
Db	602	CATTTCATGTGTGAT	614

[illegible]

Eukaryota; Viridiplantae; streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;
 Asteridae; eusterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 609)
 van der Hoeven, R., Sun, H., Bezzerides, J., Cho, J., Uterback, R.,
 Ronning, C. and Tanksley, S.
 Generation of ESTs from Tomato Suspension Cultures
 Unpublished (2001)
 Contact: CUGI

FEATURES
source location/qualifiers
1. .609

```

/organism="Lycopersicon esculentum"
/cultivar="PA496, E6203"
/db_xref="taxon:4081"
/clone="CT053N9"
/clone_1fb="ct05"
/tissue_type="suspension cultures"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
XhoI; Suspension cultures of L.esculentum E6203 were grown
in Murashige and Skoog based medium, supplemented with 15
coconut milk (filter sterilized and added after
autoclaving), 2% saccharose, and 1mg/ml 2,4d (pH5.8).
Fresh medium was added every 7 days, and cultures were
grown at 25 C, with 12hrs of light and continuous
shaking."

```

	Query Match	14.3%	Score 164.6	DB 10	Length 609
	Best Local Similarity	64.5%	Pred: No. 1.8e-29		
	Matches 273	Conservative	0	Mismatches 129	Indels 21
				Gaps	1
Qy	705	ggaatgggaactttggaatgggaagatatcagaagatggaattcaactcaatgtygat	764		
Db	24	ggaaagaggttttgaatattgagcaagattaccgaaagatggacttcactcctaaaggttaat	83		

OY	765	ttctgaagcactacgaatgtagtaataatgcgaagctgtacatagaanaagctacacga	824
Db	84	AAATTAAGCTTTAAATATGGGAATGTTGTCAAAATTTGTGCACCGAAGGCACCTTAAAGA	143
OY	825	tattaaatccaagaatatgtagtaccacaaggggtggaaggtgttcccaacttcacagc	884
Db	144	TGTCAATTTTAAAGTTATGTGATTCACGCGGGTTGGAAGGTCCTACCAATGATTCATGTC	203
OY	885	agtaacacttgatccctctctcatgaaatcccttgaatttaaccatgagatgg--	942
Db	204	TGTTCAATTTGGACCATCATGTTCAATCAATGACCTCCACTTAATCCCTGGAGATGGGA	263
OY	943	-----accaaaagcagcgcttggagagagaaagagatgctcc	983
Db	264	GAGTGATGACCAATTAAGCAGACAGACTTGACTCTTTGGGGGAGATCAAGATGTTGTCC	323
OY	984	tgtgtgtgaacttggcaagctccaattgcttcttcctcaacactgtgctcctcta	1043
Db	324	TGGATTTTGACTTGCAGAAAGTTGAAGTGATCCCTTCTCCTTCCACACCTTGTCAGAAAATA	383
OY	1044	taggttgaanaataaagtccagatgaaatgccaactcgcagcaaccttaagtgttaagag	1103
Db	384	CAGATGGGAGGTAGAAGAGGAGAACACCACTTCCTTATCCATATGTGGAGTTCAAAAA	443
OY	1104	agg 1106	
Db	444	TGG 446	

RESULT	6
LOCUS	BI204446
DEFINITION	678 bp mRNA linear EST 11-JUL-2001
ACCESSION	BI204446
VERSION	EST5722486
KEYWORDS	nrRNA sequence.
SOURCE	BI204446.1 GI:14682170
ORGANISM	EST.
REFERENCE	tomato.
AUTHORS	Lycopersicon esculentum
TITLE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
JOURNAL	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
COMMENT	Asteridae; eusterids I; Solanales; Solanaceae; Solanum;
	Lycopersicon.
	1 (bases 1 to 678)
	van der Hoeven,R., Sun,H., Bezzerides,J., Cho,J., Uteerback,R.,
	Romling,C. and Tanksley,S.
	Generation of ESTs from Tomato Suspension Cultures
	Unpublished (2001)
	Contact: CUGI
	Clemson University Genomics Institute
	100 Jordan Hall, Clemson, SC 29634, USA
	Email: http://www.genome.clemson.edu/orders/index.html .

FEATURES	source	Location/Qualifiers
	1. .678	
		/organism="Lycopersicon esculentum"
		/cultivar="F496, E6203"
		/db_xref="taxon:4081"
		/clone="CT05E13"
		/clone_11b="CT05"
		/tissue_type="suspension cultures"
		/lab_host="SOLR"
		/note="Vector: pBluescript SK(-): Site.1: EcoRI; Site.2: XhoI; Suspension cultures of L.esculentum E6203 were grown/
		in Murashige and Skoog based medium, supplemented with 15
		coconut milk (filter sterilized and added after autoclaving), 2% saccharose, and 1mg/ml 2,4d (pH5.8).
		Fresh medium was added every 7 days, and cultures were
		grown at 25 C, with 12hrs of light and continuous
		shaking."
BASE COUNT	226 a	108 c 134 g 210 t
ORIGIN		

High quality sequence stop: 420.

FEATURES
Source
Location/Qualifiers

1..667
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl066-1781"
/clone_1lb="Gm-cl066"
/tissue_type="Leaf and shoot tlp, salt stressed, 2 week old seedling"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+, Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed from mRNA isolated from unexpanded leaves and the shoot tips of 2 week old seedling from the cultivar Williams. The 2 week old seedlings were salt stressed in a solution of 500mM NaCl for 3 days prior to harvesting. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 194 a 155 c 141 g 177 t
ORIGIN

Query Match 13.2%; Score 152.2; DB 10; Length 667;
Best Local Similarity 57.4%; Pred. No. 1.8e-26;
Matches 274; Conservative 0; Mismatches 203; Indels 0; Gaps 0;

QY 1 atggatgagccttctatgagaataattcttcttaacctatgatgaagctcc 60
DB 128 ATGGGTGGCCATTTCTTGTGTGAACCATGGCTATTGAAGCTTATCTGGCCACCA 187
QY 61 atcggtacattcttgaacaacggtgttcaagtgatgagaagtggttcaagtaata 120
DB 188 ATAGGGGAATTTCATGAGACACATAGCAGGTATGTCATTTACAGTCAAAACGT 247
QY 121 tctgtggaagaacgagtagtctcatgtgacaaagtaacatgttctacttcaaac 180
DB 248 TTTGGGAGCCACATAGTGTACAGCATGACGACGACGACGACGACGACGACGAC 307
QY 181 gaaggaagatgttttccatcgattatccaaagcgatgcagacttctcgcaatat 240
DB 308 GAAGGGAAATGTTCGAGTCAGCATCTAGAACGATCGGTGAATCTAGGAAATGG 367
QY 241 tccctctatagccacgagaataatcacaagaataaataatgttattatgaact 300
DB 368 TCCATGTGTGCTTACTGTGATGATGATGATGATGATGATGATGATGATGATGAT 427
QY 301 atcaatcacaagaatgcgaactgtcttctactgagcagaagactctctatctcg 360
DB 428 CTAAAGCCAGCGCTGAGAACACACTCTTGAAGAGGTGAGGAAAGCTATCCCTTG 487
QY 361 atactaagaatcagaagaattgcccgaagtcgaattcacaagaagaatgaatgtt 420
DB 488 GTTCTGAATCTTGGAGCCAGATTCATATCTTCAGCCCAAGATGAAGTAAAGATTC 547
QY 421 actctcagtgattatgtaaaccaactcttgagatcaagccagaagaaga 477
DB 548 ACCTTCAATGTATGTGCTAAGCATATCATGAGATGATGATGATGATGATGATGATG 604

RESULT 9
BI204438 621 bp mRNA linear EST 11-Jul-2001
LOCUS BI204438
DEFINITION EST522478 cTOS lycopersicon esculentum cDNA clone cTOS5C19 5' end,
ACCESSION BI204438
VERSION BI204438.1 GI:14682162

KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 621)

REFERENCE
AUTHORS van der Hoeven,R., Sun,H., Bezzerides,J., Cho,J., Utterback,R.,
Rommung,C. and Tanksley,S.
TITLE Generation of ESTs from Tomato Suspension Cultures
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.

FEATURES
source
Location/Qualifiers

1..621
/organism="Lycopersicon esculentum"
/cultivar="TMA96, E6203"
/db_xref="taxon:4081"
/clone="cTOS5C19"
/clone_1lb="cTOS"
/tissue_type="suspension cultures"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Suspension cultures of L.esculentum E6203 were grown in Murashige and Skoog based medium, supplemented with 15% coconut milk (filter sterilized and added after autoclaving), 2% saccharose, and 1mg/ml 2,4d (pH5.8). Fresh medium was added every 7 days, and cultures were grown at 25 C, with 12hrs of light and continuous shaking."

BASE COUNT 201 a 101 c 118 g 201 t
ORIGIN

Query Match 12.5%; Score 143.4; DB 10; Length 621;
Best Local Similarity 63.5%; Pred. No. 2.5e-24;
Matches 247; Conservative 0; Mismatches 121; Indels 21; Gaps 1;

QY 739 aagatgaattcaatcgaatgtgattcttgagagcactagatggttaataatcgtaag 798
DB 1 AAGATGAGCTTCACGCAAAAGTAATTAATGAAGCTTTAAGATGGAATGTTGCAAA 60
QY 799 actgtacataagaagaactctatctgatataatcagaagaatgtgattccaaaggg 858
DB 61 TTTGTGCACCGGAAGGCACTTAAGATGTCAAAATTTAAAGATTATGATTCACGCGGT 120
QY 859 ttgaagtggttccaatcttcacagcagatcattcgatccctctctcatgaatactct 918
DB 121 TGGAAAGTCTTACCACTATTTTCAAGTGTTCATTGGACCACTTCTCATATGCA 180
QY 919 ttggaatttaatccatgagatgg-----accaaacgacgagcg 957
DB 181 CTCACCTTAATTCCTTGGAGATGGAGAGATGAGCAATAAAGCAAGAGTTGACTCCT 240
QY 958 ttggagagggagtaagggatgctctgtgtggaacttggcgaagtcacaatgttttc 1017
DB 241 TTTGGGGAGGATCAAGATGTGTCTGTGAATTTGAACCTTGAAGGTTGAAGTAGCTTTC 300
QY 1018 ttctctcaatcttgcctctctatagtgtaagaataaagtcagatgaatgccaatc 1077
DB 301 TTCCTTCACCACTTGTACAAAATATACAGATGGGAGGTGAAGAAAGAACCAACCAT 360
QY 1078 gcgcacccttaagtgagatttaagagagg 1106
DB 361 GCTTATCCATATGTGAGTCAAAAATGG 389

RESULT 10
BG316131

LOCUS	BG316131	522 bp	mRNA	linear	EST 28-NOV-2001
DEFINITION	saab76f08.y1 Gm-cl032 glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl032-2967 5' similar to TR:064989 064989 STEROID 22-ALPHA-HYDROXYLASE. [1] ; mRNA sequence.				
ACCESSION	BG316131				
VERSION	BG316131.1	GI:13125561			
KEYWORDS	EST.				
SOURCE	soybean.				
ORGANISM	Glycine max				
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.				
AUTHORS	1 (bases 1 to 522)				
TITLE	Shoemaker,R., Kelm,P., Vockin,L., Erpelting,J., Corvelli,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Knusaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepteck,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Hareey,N., Schurk,R., Ritter,E., Kohn,S., Shio,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.				
JOURNAL	Public Soybean EST Project				
COMMENT	Unpublished (1999) Contact: Shoemaker R/Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel.: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu This clone is available through: ResGen, Invitrogen Corp, 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccluresgen.com High quality sequence stop: 388. Location/Qualifiers 1..522 /organism="Glycine max" /db_xref="taxon:3847" /clone="GENOME SYSTEMS CLONE ID: Gm-cl032-2967" /clone_id="Gm-cl032" /tissue_type="Cotyledons of 8-day-old 'Williams' seedlings" /lab_host="DH10B" /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from cotyledons of 8-day-old 'Williams' seedlings which were propagated on paper towels with distilled water for 3 days (retiolated), then greenhouse grown for 5 days in potting soil. The cotyledons were flash-frozen in liquid nitrogen. Stratagene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (VA, C, or G) was added to the 3' end of the primer (GAGACGAGAGAGAGAGACTGTCGAC(TT)18V) to anchor the primer at the 5' end of the poly(A) tract. After second strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (40U/u1); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using Glycerol 1% agarose gels. cDNA Size fractionation column. The column eluent was then precipitated, redissolved, and ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+)) vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 100% of the white and 87.5% of the blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=28 and 8				

[illegible]

Tue Jul 30 08:40:09 2002

us-09-995-917a-3.rst

Page 11

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 21:28:34 ; Search time 50.3 Seconds
(without alignments)
5610.992 Million cell updates/sec

Title: US-09-995-917A-3
Perfect score: 1149
Sequence: 1 atggagatgagcttcattatg.....caaaattccttgagagatg 1149

Scoring table: IDENTITY_NIC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/prodata/2/1na/5a_COMB.seq:*
2: /cgn2_6/prodata/2/1na/5b_COMB.seq:*
3: /cgn2_6/prodata/2/1na/5a_COMB.seq:*
4: /cgn2_6/prodata/2/1na/5b_COMB.seq:*
5: /cgn2_6/prodata/2/1na/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/2/1na/backfiles1.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	97.8	8.5	1608	2	US-08-622-166A-1
2	80	7.0	4937	2	US-08-622-166A-3
3	58.6	5.1	7218	1	US-08-232-463-14
4	38.8	3.4	2694	1	US-08-465-995A-1
5	38.8	3.4	2694	2	US-08-465-994C-1
6	38.8	3.4	2694	2	US-08-101-593-1
7	38.8	3.4	2694	2	US-08-101-593-1
8	37.4	3.3	3892	4	US-08-569-214-1
9	37.4	3.3	3892	4	US-08-937-236-1
10	37.2	3.2	2694	1	US-08-465-995A-3
11	37.2	3.2	2694	2	US-08-465-994C-3
12	37.2	3.2	2694	2	US-08-966-145-3
13	37.2	3.2	2694	2	US-08-101-593-3
14	37.2	3.2	2694	1	US-08-323-170B-1
15	37	3.2	9636	4	US-08-854-441-1
16	36.8	3.2	3308	3	US-08-714-918-68
17	36.8	3.2	3308	4	US-09-265-315-68
18	36.8	3.2	3308	4	US-09-265-315-68
19	36.8	3.2	3308	4	US-09-266-417-68
20	36.2	3.2	1850	3	US-08-724-466B-3
21	36.2	3.2	1850	4	US-08-882-164D-3
22	36.2	3.2	29793	4	US-09-302-812-38
23	36.2	3.2	29793	4	US-09-311-477-38
24	35.8	3.1	4090	4	US-08-569-214-4
25	35.8	3.1	4090	4	US-08-937-236-4
26	35.2	3.1	1725	4	US-08-882-164D-31
27	35.2	3.1	2439	4	US-09-386-493-1

28	34.6	3.0	1659	1	US-08-231-729B-1	Sequence 1, Appl1
29	34.4	3.0	2277	1	US-08-676-967-2	Sequence 2, Appl1
30	34.4	3.0	2277	1	US-08-676-974-2	Sequence 2, Appl1
31	34.4	3.0	2277	2	US-09-098-487-2	Sequence 2, Appl1
32	34	3.0	1659	1	US-08-231-729B-2	Sequence 2, Appl1
33	33.6	2.9	1333	1	US-08-684-862-9	Sequence 19, Appl1
34	33.2	2.9	2659	3	US-09-029-267-19	Sequence 65, Appl1
35	33.2	2.9	8157	1	US-08-358-160-65	Sequence 70, Appl1
36	33.2	2.9	8384	1	US-08-358-160-66	Sequence 273, App
37	33.2	2.9	8590	4	US-08-385-982-273	Sequence 2, Appl1
38	33	2.9	398	1	US-08-181-271A-2	Sequence 2, Appl1
39	33	2.9	2256	1	US-08-449-315-2	Sequence 2, Appl1
40	33	2.9	2256	1	US-08-449-803-2	Sequence 2, Appl1
41	33	2.9	2256	1	US-08-449-043-2	Sequence 2, Appl1
42	33	2.9	2256	1	US-08-456-265A-2	Sequence 2, Appl1
43	33	2.9	2256	1	US-08-455-416-2	Sequence 2, Appl1
44	33	2.9	2256	1	US-08-455-244-2	Sequence 2, Appl1
45	33	2.9	2256	1	US-08-455-244-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-622-166A-1
Sequence 1, Application US/08622166A
Patent No. 5952545
GENERAL INFORMATION:
APPLICANT: KONCZ, CSABA
APPLICANT: MATHUR, JAIDEEP
APPLICANT: SZEKERES, MIKLOS
APPLICANT: ALTMANN, THOMAS
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING CYTOCHROME
TITLE OF INVENTION: P450-TYPE PROTEINS INVOLVED IN THE BRASSINOSTEROID SYNTHESIS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: P.O. Box 747
City: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/622,166A
FILING DATE: 27-MAR-1996
CLASSIFICATION: B00
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 0147-0153P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1608 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: A. thaliana
IMMEDIATE SOURCE:
LIBRARY: lambda gt10
CLONE: C204


```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: ASCII DOS/TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,994C
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MARGARET A. CHURCHILL
; REGISTRATION NUMBER: 39,944
; REFERENCE/DOCKET NUMBER: 1920-305D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213/977-1001
; TELEFAX: 213/977-1003
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2694 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; LOCATION: 1..2694
; US-08-465-994C-1

```

```

Query Match          3.4%; Score 38.8; DB 2; Length 2694;
Best Local Similarity 45.2%; Pred. No. 0.12;
Matches 142; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

```

```

QY 599 acgcaattatagaagacatgaataatgaatgaagaagaatttcgtgattcgataa 658
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1310 ATGAATATATCGAGAGACAGCTCCCTAAACCAAGTGAATGTTCTTGTCTCCGAATG 1369
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 659 ttccgaatgaagatgaagaacatgcagccatagaagccaagaaggagatgggaactt 718
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1370 GATGGATGATGATGAACATCAAGAAAGTATCCCAAGAAATGCGTAAAGTATTTT 1429
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 719 tgattgggaagattatcgaagaatgaattcctcaatgtgtgatttcgaggaactac 778
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1430 TCCAGCGTTAAGATGAGAAAGAAATGTCGCTGAAGAAATGCGCAAGCTATTTA 1489
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 779 gatgtgtaataatcgtcaagactgtacatagaagaagctactcatgtatattaatcaag 838
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1490 AAAAGATTATATGAAGAGCGCGAGGCTGTTCACACTAAACCAAGATTGAACGATATG 1549
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 839 aatagtgattccaagggtggaagggtttccaatcttcacagcagctacatcttgatc 898
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1550 TTAAGTCACTGATGATTTCTTAAATGAACATGCAATTATGTAATCTGTTCTTAATA 1609
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 899 cctctcttcataaa 912
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1610 GTCTGATTGAAGAA 1623
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 6
US-08-966-145-1
; Sequence 1, Application US/08966145A
; Patent No. 5945312
; GENERAL INFORMATION:
; APPLICANT: Goodman, Myron F.
; APPLICANT: Reha-Krantz, Linda J.
; TITLE OF INVENTION: Synthesis of Fluorophore-Labeled DNA
; FILE REFERENCE: 1920-353D1
; CURRENT APPLICATION NUMBER: US/08/966,145A
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: US 08/632,742
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 1
; LENGTH: 2694 base pairs

```

```

; TYPE: DNA
; ORGANISM: T2 bacteriophage
; PUBLICATON INFORMATION:
; AUTHORS: Goodman, Myron F.
; AUTHORS: Reha-Krantz, Linda J.
; PATENT DOCUMENT NUMBER: US 5,660,980
; FILING DATE: 1995-06-06
; PUBLICATION DATE: 1997-08-26
; US-08-966-145-1

```

```

Query Match          3.4%; Score 38.8; DB 2; Length 2694;
Best Local Similarity 45.2%; Pred. No. 0.12;
Matches 142; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

```

```

QY 599 acgcaattatagaagacatgaataatgaatgaagaagaatttcgtgattcgataa 658
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1310 ATGAATATATCGAGAGACAGCTCCCTAAACCAAGTGAATGTTCTTGTCTCCGAATG 1369
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 659 ttccgaatgaagatgaagaacatgcagccatagaagccaagaaggagatgggaactt 718
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1370 GATGGATGATGATGAACATCAAGAAAGTATCCCAAGAAATGCGTAAAGTATTTT 1429
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 719 tgattgggaagattatcgaagaatgaattcctcaatgtgtgatttcgaggaactac 778
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1430 TCCAGCGTTAAGATGAGAAAGAAATGTCGCTGAAGAAATGCGCAAGCTATTTA 1489
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 779 gatgtgtaataatcgtcaagactgtacatagaagaagctactcatgtatattaatcaag 838
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1490 AAAAGATTATATGAAGAGCGCGAGGCTGTTCACACTAAACCAAGATTGAACGATATG 1549
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 839 aatagtgattccaagggtggaagggtttccaatcttcacagcagctacatcttgatc 898
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1550 TTAAGTCACTGATGATTTCTTAAATGAACATGCAATTATGTAATCTGTTCTTAATA 1609
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 899 cctctcttcataaa 912
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1610 GTCTGATTGAAGAA 1623
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 7
US-08-101-593-1
; Sequence 1, Application US/08101593
; Patent No. 5547859
; GENERAL INFORMATION:
; APPLICANT: Goodman, Myron F.
; APPLICANT: Reha-Krantz, Linda J.
; TITLE OF INVENTION: NEW DNA SEQUENCING ENZYMS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Robbins, Berliner & Carspm
; STREET: 201 No. 5547859th Figueroa Street, Fifth Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90012-2628
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/101,593
; FILING DATE: 19930802
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Spitals, John P.
; REGISTRATION NUMBER: 29,215
; REFERENCE/DOCKET NUMBER: 1920-305
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 977-1001
; TELEFAX: (213) 977-1003
; INFORMATION FOR SEQ ID NO: 1:

```

SEQUENCE CHARACTERISTICS:
 LENGTH: 2760 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE TYPE: DNA (genomic)
 NAME/KEY: CDS
 LOCATION: 1..2760
 US-08-101-593-1

Query Match 3.48; Score 38.8; DB 1; Length 2760;
 Best Local Similarity 45.2%; Pred. No. 0.12;
 Matches 142; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

QY 599 acgcaatcattgaaagacacgtaataatgcaataagaagaagcttctcgattcgtaa 658
 DB 1376 ATGAATATATCGCAGAACACCTCTTAACCAAGTATGATATCTTGTCTCCGATG 1435
 QY 659 ttctgaatgaaagatgaaacatgcaagcattgagccaaagaagggatcggggaacttt 718
 DB 1436 GATGAGTATGATTAACATCAAGAACGATATTCATTCCAAGGAATGCTAAAGTATTTT 1495
 QY 719 tgaattggagaattatcagaagaatggaattcactcaatgctgattctcgaggcaactac 778
 DB 1496 TCCAGCGTTAAAGATTGGAAGAAAGAAATGTTCCCTGAAGAAATGATGCCGAAGCTATTA 1555
 QY 779 gattggttaaatcgtcaagacgtacatagaaagcactactcattgatttaattcaaga 838
 DB 1556 AAAAGATTATTAAGAAAGCGGCGCTGTTGTTCACTTAACCAAGAACTTGAACGATATG 1615
 QY 839 aatatgcatcacaagaagggtggaagggtgtttccaaatctcacaagcagtaactcgtac 898
 DB 1616 TTAAGTTCACTGATGATTTCTTAATGAAGTAATGATTAATGATTAATGATTAATA 1675
 QY 899 cctctctcatgaa 912
 DB 1676 GTCGATTGAAGAA 1689

RESULT 8
 US-08-569-214-1
 Sequence 1, Application US/08569214
 Patent No. 6165469
 GENERAL INFORMATION:
 APPLICANT: MANN, BARBARA J.
 TITLE OF INVENTION: RECOMBINANT ENTAMOEBA HISTOLYTICA LECTIN
 TITLE OF INVENTION: SUBUNIT PEPTIDES AND REAGENTS SPECIFIC FOR MEMBERS OF THE
 NUMBER OF SEQUENCES: 170 KD SUBUNIT MULTIGENE FAMILY
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 2000 PENNSYLVANIA AVENUE N.W., STE. 5500
 CITY: WASHINGTON
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/569,214
 FILING DATE:
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/06890
 FILING DATE: 17-JUN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.

REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 9148-0006.21
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 TELEX: 90-4030
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3892 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: join(1..3873, 3877..3882, 3886..3891)
 US-08-569-214-1

Query Match 3.3%; Score 37.4; DB 4; Length 3892;
 Best Local Similarity 53.8%; Pred. No. 0.35;
 Matches 77; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 573 tagatcgaatcgtaatacattcaaaacgcaattatagaagacatgaaatgcaataag 632
 DB 234 TAAATATACTATCAATGAAGATCATATGTAAGTTGAAGTATTAATAACAAATAT 293
 QY 633 aagaagaagtttctcgattcgataattcgaaatgaaatgaaacacagcagcattag 692
 DB 294 TAAACAAAGTTTGTCAAAAAGAAATATGCAATTCATTAATGAAAATATGAAGTTGAT 353
 QY 693 agccaagaagggtggtgggaac 715
 DB 354 GGACAATGTTCCAGTTGATGAAC 376

RESULT 9
 US-08-937-236-1
 Sequence 1, Application US/08937236
 Patent No. 6187310
 GENERAL INFORMATION:
 APPLICANT: MANN, BARBARA J.
 APPLICANT: PETRI, WILLIAM A.
 TITLE OF INVENTION: RECOMBINANT ENTAMOEBA HISTOLYTICA LECTIN
 TITLE OF INVENTION: SUBUNIT PEPTIDES AND REAGENTS SPECIFIC FOR MEMBERS OF THE
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 2000 PENNSYLVANIA AVENUE N.W., STE. 5500
 CITY: WASHINGTON
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/937,236
 FILING DATE:
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/569,214
 FILING DATE: 16 SEPTEMBER 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: LIYNAT, SHMUEL
 REGISTRATION NUMBER: 33,949
 REFERENCE/DOCKET NUMBER: 291482000622
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763

TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3892 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: join(1..3873, 3877..3882, 3886..3891)
US-08-937-236-1

Query Match 3.3%; Score 37.4; DB 4; Length 3892;
Best Local Similarity 53.8%; Pred. No. 0.35;
Matches 77; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy 573 tagatccaatcgtaataatcatcaaaagcgaattatagaagacatgaataatgcatatag 632
Db 234 TAAATACTACTCAATGAAGATCATATGTAAAGTTGAAGATTAATTAACCAATAT 293
Qy 633 agagaagaattttctgattcgataatcgaatgaagaatgaagaacatgcacatag 692
Db 294 TAAACAGATTITTTGCAAAAAGAAATATGCATATCCAAATTGAATAATATGAAGTTGTTG 353
Qy 693 agccaagaagggatgagggac 715
Db 354 GGACAAATGTTCCAGTTGATGAAC 376

RESULT 10
US-08-465-995A-3
Sequence 3, Application US/08465995A
Patent No. 5660980

GENERAL INFORMATION:
APPLICANT: MYRON F. GOODMAN
TITLE OF INVENTION: METHODS FOR IDENTIFYING AND ISOLATING
TITLE OF INVENTION: VARIANT T4 POLYMERASES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Robbins, Berliner & Carson, LLP
STREET: 201 No. 5660980th Figueroa Street, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90012-2628
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: ASCII DOS/TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465, 995A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Robert Berliner
REGISTRATION NUMBER: 20,121
REFERENCE/DOCKET NUMBER: 1920-305D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213/977-1001
TELEFAX: 213/977-1003
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2694 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE: CDS
LOCATION: 1..2694
US-08-465-995A-3

Query Match 3.2%; Score 37.2; DB 1; Length 2694;
Best Local Similarity 44.9%; Pred. No. 0.34;
Matches 141; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

Qy 599 agcgaattatagaacatgaataatgaataaagaagaagatttctgattcgataa 658
Db 1310 ATGATATATATCGCAGGAACACTCTCTTAACCGAGATGATGATATTCCTGTCGGAATG 1369
Qy 659 ttccgaatgaagatgaagaacatgcacatcattagaagccaagaagggatggggaattc 718
Db 1370 GATGATGATATGATTAACATCAAGAGGTATCATTCACCAAGAAATGCTTAAGTATTTT 1429
Qy 719 tgaattgggaagatataagaagaatggaattcactcaatggtgatttcgagggcactac 778
Db 1430 TCCACGCTTAAGACTCGAAGAAAGAAATGTTGCTGGAAGAAATGAAATGCCAAGCTATTA 1489
Qy 779 gatggtgaataatcgtaacatgtaacataagaagaagctactcatgataatcaatcaag 838
Db 1490 AAAAGATTATATGAAAGCGCAGGCTCTGTTCAACTTAACCAAGAGTTGAACGATATG 1549
Qy 839 aatatgattccaaagggtgggaagtgttccaatctcaagcgatgatacttgatc 898
Db 1550 TTAAGTTCAGTGATGATTTCTTAATGAACATGCAATTAACCGAATCTGTTCAATA 1609
Qy 899 cctctctcatgaa 912
Db 1610 GTCGTATTGAAGA 1623

RESULT 11
US-08-465-994C-3
Sequence 3, Application US/08465994C
Patent No. 5928919

GENERAL INFORMATION:
APPLICANT: MYRON F. GOODMAN
TITLE OF INVENTION: VARIANT DNA POLYMERASES
TITLE OF INVENTION: VARIANT DNA POLYMERASES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Robbins, Berliner & Carson, LLP
STREET: 201 No. 5928919th Figueroa Street, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90012-2628
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: ASCII DOS/TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465, 994C
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MARGARET A. CHURCHILL
REGISTRATION NUMBER: 39,944
REFERENCE/DOCKET NUMBER: 1920-305D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213/977-1001
TELEFAX: 213/977-1003
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2694 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE: CDS
LOCATION: 1..2694
US-08-465-994C-3

Query Match 3.2%; Score 37.2; DB 2; Length 2694;

Best Local Similarity 44.9%; Pred. No. 0.34; Matches 141; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 599 acgcaattatagaagacatgaataatgaataagagaagaattcttcgattcgataa 658
DB 1310 ATGAAATATATCGCAGCAACAGCTCTTAACCGAGTATGTAATATTCTTGTTCGCAAG 1369
QY 659 ttccgaatgaagaatgaagaacatgcagccattgaagcgaagaaggatgggaactt 718
DB 1370 GATCGATGATGTGATTAACATCAAGAAAGTATCTTCCAAAGAAATCGCTAAATTTT 1429
QY 719 tgaattgggaagattatcagaagatgcgaattcgaatgtggaattcttgaggcaactac 778
DB 1430 TCCAGCGCTAAGACTGGAAGAAAGAAATGTTCCGTAAGAAATGATCCGAAGCTATTA 1489
QY 779 gatcgtggaatatacgcgaagacatgcataagaagaagctactcgaatataatcaag 838
DB 1490 AAAAGATTATTAAGAAAGCGGAGGCTTGTTCAACTTAACGAGAAATTGAACGATATG 1549
QY 839 aataatgattccaagaagggtggaagggtttcccaattcgaagagataacttcgac 898
DB 1550 TTAAAGTTCACTGATGATTTCTTAATGAACTATCGAATTACCGCAATCTGTTCTCAATA 1609
QY 899 cctctctcatgaa 912
DB 1610 GTCTGATTGAGAA 1623

RESULT 12

US-08-966-145-3

Sequence 3, Application US/08966145A

Patent No. 5945312

GENERAL INFORMATION:

APPLICANT: Goodman, Myron F.

APPLICANT: Reha-Krantz, Linda J.

TITLE OF INVENTION: Synthesis of Fluorophore-Labeled DNA

FILE REFERENCE: 1920-353D1

CURRENT APPLICATION NUMBER: US/08/966,145A

EARLIER FILING DATE: 1997-11-07

EARLIER FILING DATE: 1996-04-15

NUMBER OF SEQ ID NOS: 4

SEQ ID NO 3

LENGTH: 2694 base pairs

TYPE: DNA

ORGANISM: T4 bacteriophage

PUBLICATION INFORMATION:

AUTHORS: Goodman, Myron F.

AUTHORS: Reha-Krantz, Linda J.

PATENT DOCUMENT NUMBER: US 5,660,980

PATENT FILING DATE: 1995-06-06

PUBLICATION DATE: 1997-08-26

US-08-966-145-3

Query Match 3.2%; Score 37.2; DB 2; Length 2694;

Best Local Similarity 44.9%; Pred. No. 0.34; Matches 141; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 599 acgcaattatagaagacatgaataatgaataagaagaagaatttcttgattcgataa 658
DB 1310 atgaattatcgcaggaacagctcctaaacgagtgaatattcttcttcctcgaatg 1369
QY 659 ttccgaatgaagaatgaagaacatgcagccattagaagccaagaaggatgggaactt 718
DB 1370 gatcgtggaatatacgcgaagacatgcataagaagaaggatgcgaagcaatctt 1429
QY 719 tgaattgggaagattatcagaagatgcgaattcgaatgtggaattcttcgaaggcaactac 778
DB 1430 tccagctaaagaactcgaagaagaagaatgttcgctgaagaagaatgaatgcggaagcattata 1489

QY 779 gatcgtggaatatacgtcgaagactgtacataagaagaagctactcgaatataatccaag 838
DB 1490 aaaaagattatagaaggcgagggctgttccaactaaacagaagtgaagacatatg 1549
QY 839 aataatgattccaagaagggtggaagggtttcccaattcctcgaagcaatcttcgac 898
DB 1550 ttaagtcagtgatattcttcaatgaactcgaatatacgaacgaactcgttccaata 1609
QY 899 cctctctcatgaa 912
DB 1610 gtcgattgaagaa 1623

RESULT 13

US-08-101-593-3

Sequence 3, Application US/08101593

Patent No. 5547859

GENERAL INFORMATION:

APPLICANT: Goodman, Myron F.

APPLICANT: Reha-Krantz, Linda J.

TITLE OF INVENTION: NEW DNA SEQUENCING ENZYMES

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESS: 201 No. 5547859th Figueroa Street, Fifth Floor

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90012-2628

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/101,593

FILING DATE: 19930802

CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:

NAME: Spitals, John P.

REGISTRATION NUMBER: 29,215

REFERENCE/DOCKET NUMBER: 1920-305

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 977-1001

TELEFAX: (213) 977-1003

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 2760 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 1..2760

US-08-101-593-3

Query Match 3.2%; Score 37.2; DB 1; Length 2760;

Best Local Similarity 44.9%; Pred. No. 0.35; Matches 141; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 599 acgcaattatagaagacatgaataatgaataagaagaagaatttcttgattcgataa 658
DB 1376 ATGAATATATCGCAGCAACAGCTCTTAACCGAGTATGTAATATTCTTGTTCGCAAG 1435
QY 659 ttccgaatgaagaatgaagaacatgcagccattagaagccaagaaggatgggaactt 718
DB 1436 GATCGATGATGTGATTAACATCAAGAAAGTATCTTCCAAAGAAATGCTAAATTTT 1495
QY 719 tgaattgggaagattatcagaagatgcgaattcgaatgtggaattcttcgaaggcaactac 778

Tue Jul 30 08:40:09 2002

us-09-995-917a-3.rni

Page 9

Db 5102 GATGACACATTAGAGAGAGTTATATATATATGATATTG 5138

Search completed: July 29, 2002, 22:40:16
Job time: 4302 sec

THIS PAGE BLANK (USPTO)